GENÉTICA
Essencial
Márcia Pimentel | Cíntia Santos-Rebouças | Cláudia Gallo
Organizadores: Carlos Alberto Mourão Júnior & Dimitri Marques Abramov
gen
GUANABARA
KOOGAN

GENÉTICA
Essencial

Material Suplementar

Este livro conta com o seguinte material suplementar:

- Ilustrações da obra em formato de apresentação (acesso restrito a docentes)

Como usar

O acesso ao material suplementar é gratuito. Basta que o leitor se cadastre e faça seu *login* em nosso site (www.grupogen.com.br), clicando no *menu* superior do lado direito e, após, clique em *GEN-IO*.

É rápido e fácil. Caso haja alguma mudança no sistema ou dificuldade de acesso, entre em contato conosco (sac@grupogen.com.br).

GEN-IO (GEN | Informação Online) é o ambiente virtual de aprendizagem do GEN | Grupo Editorial Nacional, maior conglomerado brasileiro de editoras do ramo científico-técnico-profissional, composto por Guanabara Koogan, Santos, Roca, AC Farmacêutica, Forense, Método, Atlas, LTC, E.P.U. e Forense Universitária.

Os materiais suplementares ficam disponíveis para acesso durante a vigência das edições atuais dos livros a que eles correspondem.

Autoras

Márcia Mattos Gonçalves Pimentel

Bacharel em Ciências Biológicas. Mestre em Ciências Biológicas (Genética Humana) pela Universidade Federal do Rio de Janeiro (UFRJ). Doutora em Ciências (Genética Humana) pela UFRJ. Professora Associada do Departamento de Genética do Instituto de Biologia Roberto Alcântara Gomes da Universidade do Estado do Rio de Janeiro (UERJ). Coordenadora do Serviço de Genética Humana da UERJ.

Cláudia Vitória de Moura Gallo

Bacharel em Ciências Biológicas, modalidade Médica pela Universidade do Estado do Rio de Janeiro (UERJ). Mestre em Ciências (Bioquímica) pela Universidade Federal do Rio de Janeiro (UFRJ). Doutora em Biologia Molecular pelo Instituto Jacques Monod, Universidade Paris VII, Paris – França. Professora Associada do Departamento de Genética do Instituto de Biologia Roberto Alcântara Gomes da UERJ.

Cíntia Barros Santos-Rebouças

Bacharel em Ciências Biológicas pela Universidade do Estado do Rio de Janeiro (UERJ). Doutora em Ciências (Genética Humana) pela UERJ. Professora Adjunta do Departamento de Genética do Instituto de Biologia Roberto Alcântara Gomes da UERJ. Coordenadora Adjunta do Serviço de Genética Humana da UERJ.

Organizadores da série Essencial

Carlos Alberto Mourão Júnior

Médico Endocrinologista. Matemático. Mestre em Ciências Biológicas pela Universidade Federal de Juiz de Fora. Doutor em Ciências pela Escola Paulista de Medicina – UNIFESP. Pós-Graduado em Filosofia pela UFJF. Professor Adjunto de Biofísica e Fisiologia da Universidade Federal de Juiz de Fora.

Dimitri Marques Abramov

Médico Psiquiatra. Mestre em Ciências Biológicas pela Universidade Federal de Juiz de Fora. Doutor em Ciências pelo Instituto de Biofísica Carlos Chagas Filho – UFRJ.

EDITORA GUANABARA KOOGAN LTDA.
Uma editora integrante do GEN | Grupo Editorial Nacional
Travessa do Ouvidor, 11
Rio de Janeiro – RJ – CEP 20040-040
Tels.: (21) 3543-0770/(11) 5080-0770 | Fax: (21) 3543-0896
www.editoraguanabara.com.br | www.grupogen.com.br | editorial.saude@grupogen.com.br

■ Capa: Bruno Sales

Editoração eletrônica: Anthares

Projeto gráfico: Editora Guanabara Koogan

■ Ficha catalográfica

P699g

Pimentel, Márcia Mattos Gonçalves
Genética essencial / Márcia Mattos Gonçalves Pimentel, Cláudia Vitória de Moura Gallo, Cíntia Barros Santos-Rebouças . - [Reimpr.]. – Rio de Janeiro : Guanabara Koogan, 2017.

ISBN 978-85-277-2189-9

1. Genética humana. 2. Genética médica. I. Gallo, Cláudia Vitória de Moura. II. Santos-Rebouças, Cíntia Barros. III. Título.

12-8185.

CDD: 616.042
CDU: 616-056.7

Dedicatória

A meu esposo, Carlos, grande parceiro, pela compreensão e pela cumplicidade de mais de 30 anos. A Júlio Cezar e Augusto Cezar (filhos queridos) e a Tássia ("nora-filha" de coração), pelo carinho e pelo incentivo. Aos familiares e amigos, por acreditarem e apoiarem.

Márcia Mattos Gonçalves Pimentel

A minha família e amigos, pelo apoio incondicional.

Cláudia Vitória de Moura Gallo

Aos meus progenitores, Antonio e Olga, por terem semeado em mim a paixão pelo aprendizado. A Airton, por compreender minhas ausências, e ao resultado de nossa mais bela recombinação meiótica, nossa princesa Beatriz. À Mercedes, pelo incentivo de sempre e à família Rebouças pelo apoio.

Cíntia Barros Santos-Rebouças

Agradecimentos

Agradecemos à autora Márcia Pimentel, pela liderança editorial e pela oportunidade de participar deste aprazível desafio.

Cíntia Barros Santos-Rebouças e
Cláudia Vitória de Moura Gallo

Agradecemos a todos os profissionais da Editora Guanabara Koogan que contribuíram de maneira direta ou indireta para a concretização desta obra.

Expressamos, ainda, nosso reconhecimento especial ao Dr. Carlos Alberto Mourão Júnior (Professor Adjunto de Biofísica e Fisiologia da Universidade Federal de Juiz de Fora), cuja experiência em obras anteriores e ideias inspiradoras foram fundamentais para o aprimoramento deste projeto.

Márcia Mattos Gonçalves Pimentel,
Cíntia Barros Santos-Rebouças e
Cláudia Vitória de Moura Gallo

Prefácio

Só existe um meio de multiplicar o conhecimento: repartindo-o. Foi essa certeza que nos motivou a aceitar o desafio de passar para o papel um pouco de nossa experiência no ensino da Genética, ciência que, continuamente, renova-se em uma velocidade espantosa.

Em 1906, William Bateson, ao introduzir o termo "Genética" como uma nova subárea da Fisiologia durante o Congresso de Botânica em Londres, não podia imaginar os desmembramentos que isso teria. O século 20, registrado por Evelyn Keller como o "século do gene", contabiliza inúmeros avanços que ampliaram nossa visão dessa ciência. As bases da Genética Molecular emergiram em 1953 com a descoberta da estrutura do DNA, alcançaram a tecnologia do DNA recombinante em meados dos anos 1970 e culminaram com o sequenciamento do genoma humano no início do século atual. Trata-se de uma das áreas de conhecimento que mais se desenvolveram nas últimas décadas e cujos avanços revelam efeitos surpreendentes, com amplas aplicações em diversas áreas. Assim, para todos os que buscam utilizar e se beneficiar desses avanços, o primeiro e decisivo passo é conhecer a Genética em seus aspectos básicos.

Com esse intuito, organizamos esse livro que, em sua primeira edição, faz uma incursão por todos os conceitos da Genética, essenciais à formação de um profissional da área biomédica. Esperamos que essa obra desperte no leitor o crescente interesse pela Genética.

Composição do texto e público-alvo

Nesta obra, almejamos alcançar dois públicos-alvos: alunos que estão iniciando seu aprendizado sobre Genética e docentes que ministram aulas nessa área de conhecimento. Para atendermos aos discentes, tivemos o cuidado de fazer uma abordagem clara, de leitura aprazível, utilizando figuras que ilustrassem o conteúdo de maneira objetiva. Além disso, em alguns momentos, suprimimos aspectos e discussões mais aprofundadas e buscamos democratizar o conhecimento, pontuando sua aplicação no cotidiano do leitor. Acreditamos que este formato "Essencial" possibilita ao aluno assimilar as informações complexas pertinentes à área de Genética de modo autoexplicativo, o que propicia um bom aproveitamento da aprendizagem em sala de aula. Paralelamente, com a seleção dos conteúdos abordados, almejamos, ainda, abranger todas as etapas da prática docente, incluindo o planejamento das atividades didáticas, a atualização do professor e a utilização de modelos para avaliação do conhecimento adquirido pelo aluno. Para facilitar a interpretação e o encadeamento dos conceitos explorados, alguns capítulos podem demandar mais de uma aula, enquanto outros podem ser explorados em um período menor.

Organização do texto

Este livro reúne os principais assuntos da área. Nos três capítulos iniciais, focamos a Genética Molecular, detalhando como o material genético se organiza dentro das células, como a informação biológica é perpetuada e expressa pelos genes e como as alterações nesses processos, quando não reparadas, podem levar ao desenvolvimento de doenças.

Nos Capítulos 4, 5 e 6, destacamos como os cromossomos se dividem e como as alterações no processo de divisão celular podem resultar em fenótipos anormais. Nesses capítulos exploramos, ainda, os diferentes tipos de herança existentes. Já no Capítulo 7, descrevemos conceitos relacionados com a Genética Bioquímica e a Imunogenética e, no Capítulo 8, mostramos como ocorre a dinâmica dos genes nas populações, explorando os mecanismos pelos quais as forças evolutivas atuam. Por fim, nos Capítulos 9 e 10, utilizamos o câncer como doença modelo, a fim de estudarmos conceitos adicionais de Genética e descrevemos as possíveis aplicações de metodologias moleculares.

Estrutura dos capítulos

O leitor encontrará algumas seções específicas que irão ajudá-lo a assimilar o conteúdo estudado. No início dos capítulos é apresentada uma lista de conceitos-chave essenciais ao tema, imprescindíveis para a compreensão do texto. Ao final dos capítulos, são apresentados um pequeno resumo, ilustrando as principais ideias abordadas, e algumas questões de autoavaliação. Sugerimos que essas questões sejam trabalhadas pelo professor ao fim de cada aula ministrada, para estimular o aluno a buscar respostas adequadas.

Destaques

É importante que os leitores discentes e docentes fiquem atentos aos destaques distribuídos ao longo do texto da obra. Os trechos em azul apresentam pontos fundamentais para a interpretação do texto ou conceitos consolidados na Genética. Já os destaques em ocre representam a definição de termos genéticos específicos empregados na obra, com os quais o estudante do primeiro ano ainda não está familiarizado. Esses termos estão todos reunidos, em ordem alfabética, no Glossário ao final do livro, a fim de facilitar a consulta. Curiosidades e/ou atualidades dos capítulos também podem ser encontradas nos boxes Genética em Foco.

Bibliografia

Procuramos utilizar como referência livros-texto e artigos científicos atuais e de excelência, mas, tendo em vista que a Genética é uma ciência cujo conhecimento evolui a cada dia, eventualmente algumas informações podem se tornar desatualizadas.

Boa leitura!

Sumário

GENÉTICA

Essencial

Como usar as características especiais deste livro

- Termos fundamentais são destacados no texto e definidos nas margens. Esse recurso evita que a leitura seja interrompida e serve de elemento de revisão dos assuntos. Essas palavras estão repetidas no Glossário, ao final do livro

- Os boxes Genética em Foco descrevem curiosidades práticas acerca da Genética

- Destaques em azul consolidam conceitos descritos no texto

- Todos os capítulos se iniciam com o item Objetivos de estudo, que relaciona os principais aspectos que devem ser compreendidos ao término da leitura

- Relação de Conceitos-chave do capítulo, fundamentais para a compreensão da Genética

- A Introdução do texto principal dos capítulos contém uma visão geral daquilo que será abordado em seguida

- O Resumo ao final de cada capítulo possibilita revisões rápidas do texto, além de ser uma ferramenta útil na preparação para testes e provas

- Perguntas de Autoavaliação possibilitam a aferição dos conhecimentos adquiridos

1

O Material Genético | Composição, Estrutura, Função e Duplicação

Objetivos de estudo

Saber onde o material genético está armazenado nas células

Conhecer a composição química e a estrutura molecular dos ácidos nucleicos (DNA e RNA), bem como as diferentes conformações assumidas pela dupla hélice de DNA

Conhecer os tipos de sequências de DNA que formam o genoma

Conhecer a constituição e a organização gênica no genoma humano

Compreender como a cromatina se organiza no núcleo eucariótico, com diferentes níveis de condensação

Reconhecer os eventos que ocorrem nas diferentes fases do ciclo celular

Entender como se processa a duplicação do DNA em procariotos e eucariotos, reconhecendo a função dos inúmeros fatores proteico-enzimáticos envolvidos no processo

Compreender a função dos telômeros e a ação da telomerase na síntese das extremidades dos cromossomos lineares eucarióticos

Conceitos-chave do capítulo

Ácido desoxirribonucleico (DNA)
Ácido nucleico
Ácido ribonucleico (RNA)
Autossomo
Base nitrogenada
Bolha de replicação
Célula somática
Centrômero
Ciclo celular
Citocinese
Cloroplasto
Cromátide
Cromossomo
Cromossomo dicêntrico
Cromossomo em anel
Cromossomo unifilamentar (uninêmico)
Desoxirribose
Dextrógira
Diploide
DNA
DNA helicase
DNA ligase
DNA polimerase
DNA repetitivo
DNA satélite
DNA topoisomerase
Dupla hélice
Duplicação semiconservativa
Elemento de transposição
Encadeamento (*splicing*)
Endonuclease
Espermatozoide
Eucarioto
Fita atrasada ou tardia (*lagging strand*)
Fita líder (*leading strand*)
Forquilha de replicação
Fragmento de Okazaki
Gene
Gene ortólogo
Genoma
Heterocromatina
Histona
Interfase
Levógira
Megabase (Mb)
Meiose
Minissatélite
Microssatélite
Mitocôndria
Mitose
Nucleossomo
Nucleosídio
Nucleotídio
Origem de replicação
Ovócito
Pentose
Pequeno RNA nuclear (snRNA)
Pequeno RNA nucleolar (snoRNA)
Pirimidina
Polipeptídio
Poliploide
Ponte de hidrogênio
Primase
Primer (oligonucleotídio iniciador)
Procarioto
Proteína de ligação ao DNA fita simples (SSB)
Purina
Quilobase (kb)
Região subtelomérica
Réplicon (unidade de replicação)
Replissomo
Retrotransposon
Retrovírus
Ribose
RNA
RNA mensageiro (mRNA)
RNA ribossômico (rRNA)
RNA transportador (tRNA)
Solenoide
Telomerase
Telômero
Transcrição reversa
Zigoto

Introdução

Todos os seres vivos, nas suas mais diferentes formas, são constituídos por uma extensa diversidade de células, e são elas que contêm o conjunto de informações vitais à sua própria sustentação, precisamente, a informação genética, tanto em organismos unicelulares, como as bactérias, quanto nos demais seres, incluindo o homem. As células são capazes de armazenar e decodificar essas informações de maneira idêntica, utilizando os mesmos códigos e sistemas de processamento, de maneira que a maioria das funções celulares básicas é realizada por meio de processos bioquímicos muito semelhantes nos diferentes organismos. Esse complexo programa de desenvolvimento celular, geneticamente determinado e inerente a todo ser vivo, é preservado em sua integridade e capaz de ser transmitido, pelas divisões celulares, de uma célula às suas células-filhas.

Os primeiros conhecimentos que levariam ao entendimento da base bioquímica da hereditariedade e da variabilidade genética surgiram no início do século XX, porém, somente em 1944, Avery, MacLeod e McCarty demonstraram que a informação genética nas células estava contida no ácido desoxirribonucleico (DNA).

Os estudos sobre a natureza química desta molécula progrediram até que sua estrutura fosse enfim definida, em 1953, pelo biólogo norte-americano James Watson e o físico inglês Francis Crick. Eles propuseram que a estrutura tridimensional do DNA seria de uma *dupla hélice helicoidal*, com duas cadeias complementares de orientações opostas, constituídas por três componentes químicos: *açúcar*, *fosfato* e *bases nitrogenadas*. Resíduos de açúcar e monofosfato alternados dariam sustentação à molécula, formando um esqueleto, e, na parte interna, as bases nitrogenadas estariam dispostas e pareadas. Essa descoberta foi um importante marco na ciência, fundamental para a elucidação dos mecanismos moleculares da hereditariedade, representando a transição da genética clássica para a genética moderna. A segunda metade do século XX foi marcada por grandes avanços metodológicos e tecnológicos, em paralelo com inúmeras descobertas relativas à genética humana, que culminaram com o mapeamento do genoma humano no início do século XXI.

O material genético

Em cada uma das cerca de 65 trilhões de células que formam um ser humano adulto de estatura média, assim como na maioria dos organismos, a informação genética está armazenada em macromoléculas de ácido desoxirribonucleico (DNA), e os seres vivos podem ser classificados como eucariotos (plantas, animais, fungos e seres humanos) ou procariotos (bactérias) com base na maneira como o DNA está disposto na célula – se em um compartimento delimitado por uma membrana (*núcleo*) ou livre no interior da célula. Nas células eucarióticas, a maior parte da informação genética está armazenada no núcleo, embora, além do DNA nuclear, também exista DNA nas mitocôndrias e nos cloroplastos dos vegetais e das algas. Cabe ressaltar que em alguns vírus a informação genética não está presente no DNA e sim no ácido ribonucleico (RNA). Este é o caso de um retrovírus muito conhecido, denominado HIV I, que causa a AIDS.

O DNA é o sítio de expressão gênica, que determina todas as funções celulares necessárias para o desenvolvimento de um organismo, de maneira que, algumas características, intrínsecas à molécula, possibilitam sua autoduplicação de modo preciso, garantindo-lhe a capacidade de transmitir a informação genética através das gerações.

Eucarioto
Organismo cujas células têm núcleo verdadeiro, ou seja, delimitado por uma membrana nuclear

Procarioto
Organismo unicelular, como as bactérias, que não tem núcleo delimitado por uma membrana

Mitocôndria
Organela citoplasmática das células eucarióticas na qual se processa a respiração celular. Ela contém DNA e se autoduplica

Cloroplasto
Organela citoplasmática de plantas e algas, que contém DNA e clorofila. No cloroplasto, ocorre a fotossíntese; também se autoduplica

Composição e estrutura dos ácidos nucleicos

Ácido nucleico
Macromolécula formada por subunidades denominadas nucleotídios que armazena a informação genética. Existem dois tipos de ácidos nucleicos: DNA e RNA

Nucleotídio
Subunidade que forma os ácidos nucleicos, constituída por um grupo fosfato, uma pentose e uma base nitrogenada

Pentose
Molécula de açúcar (desoxirribose no DNA e ribose no RNA) com cinco átomos de carbono presente nos ácidos nucleicos

Base nitrogenada
Molécula heterocíclica formada por átomos de carbono e de nitrogênio. Existem dois tipos de bases nitrogenadas: pirimidinas (T, C e U) e purinas (A e G)

Desoxirribose
Molécula de açúcar com cinco átomos de carbono que forma o DNA

Ribose
Molécula de açúcar com cinco átomos de carbono que forma o RNA

Os ácidos nucleicos (DNA e RNA) são longos polímeros formados por subunidades denominadas *nucleotídios*. Cada nucleotídio é formado pela união de três moléculas: um açúcar com cinco átomos de carbono denominado pentose, um grupo *fosfato* e uma base nitrogenada (Figura 1.1).

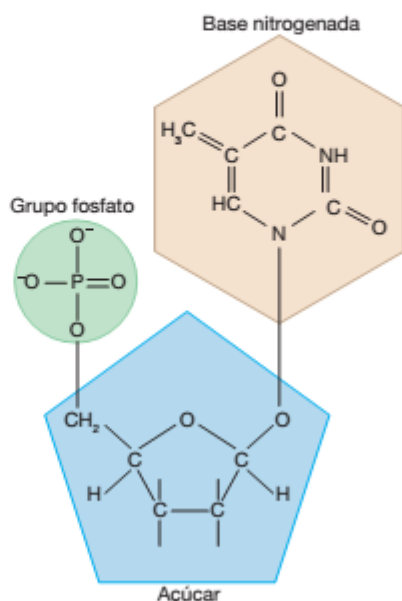



Figura 1.1 Estrutura química do nucleotídio: pentose + grupo fosfato + base nitrogenada.

No ácido desoxirribonucleico, o açúcar é a desoxirribose, e no ácido ribonucleico é a ribose. Os átomos de carbono da desoxirribose ou da ribose são numerados de 1 a 5, acrescidos do sinal de apóstrofe (Figura 1.2A). As bases nitrogenadas são moléculas heterocíclicas compostas por

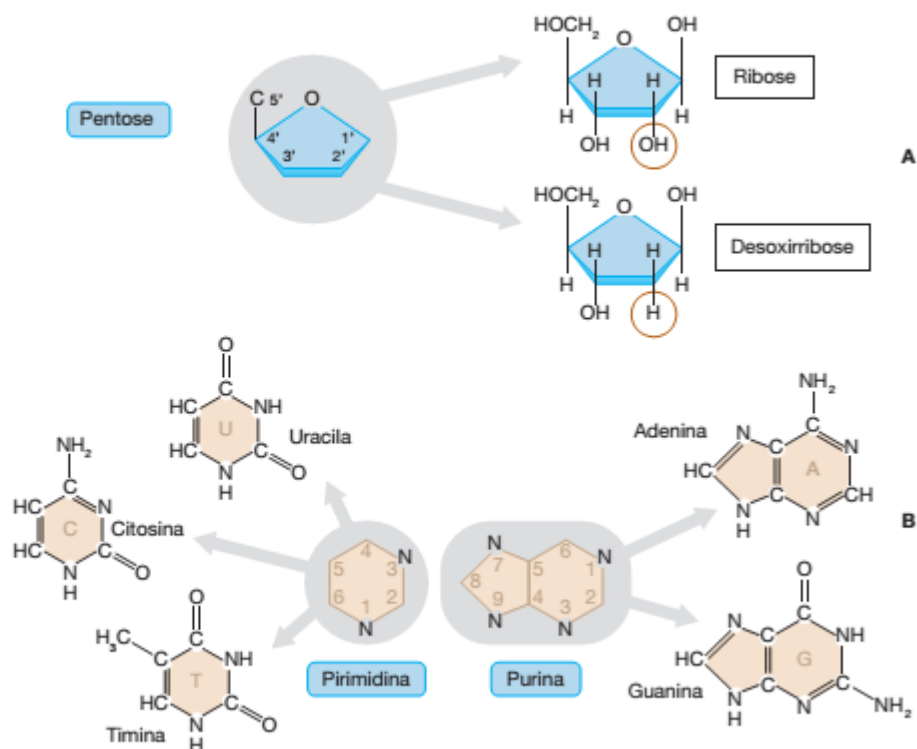



Figura 1.2 Componentes químicos dos nucleotídios. (A) Estrutura química da pentose: açúcar com cinco carbonos numerados de 1′ a 5′ ("carbono 1 linha" etc.). A diferença entre o açúcar ribose (presente no RNA) e a desoxirribose (presente no DNA) está no fato de a desoxirribose apresentar um átomo de H no carbono 2′, em vez do grupo hidroxila (OH). (B) Estrutura química das bases nitrogenadas: purinas e pirimidinas.

Pirimidina
Base nitrogenada formada por um anel heterocíclico presente nos ácidos nucleicos (C e T no DNA, e U no RNA)

Purina
Base nitrogenada formada por dois anéis heterocíclicos presente nos ácidos nucleicos (A e G no DNA e no RNA)

Nucleosídio
Subunidade que forma os ácidos nucleicos constituída pela ligação de uma base nitrogenada a uma molécula de açúcar

anéis que contêm átomos de carbono e de nitrogênio. Elas podem ser de dois tipos: aquelas formadas por um único anel, denominadas pirimidinas (citosina [C], timina [T] e uracila [U]), ou aquelas com dois anéis heterocíclicos ligados, denominadas purinas (adenina [A] e guanina [G]). Os átomos de carbono e nitrogênio das bases nitrogenadas são numerados de 1 a 9 nas purinas e de 1 a 6 nas pirimidinas. As mesmas bases estão presentes nas moléculas de DNA e de RNA, com exceção da timina, existente apenas no DNA, e da uracila, presente apenas no RNA (Figura 1.2B).

A formação da cadeia de polinucleotídios nas moléculas de DNA ou de RNA ocorre por *ligações covalentes fosfodiéster*. Conforme mostrado na Figura 1.3A, um grupo fosfato liga o átomo de carbono 5′ do açúcar (desoxirribose ou ribose) ao carbono 3′ do resíduo de açúcar seguinte, formando o "esqueleto" externo que é a parte invariável da molécula. A ligação da base nitrogenada ao açúcar forma o nucleosídio e ocorre pela união do átomo de nitrogênio 1 das pirimidinas ou do nitrogênio 9 das purinas ao carbono 1′ da pentose, e é denominada *ligação N-glicosídica* (Figura 1.3B). As bases nitrogenadas estão voltadas para dentro da molécula, e sua sequência ao longo do polímero constitui sua porção variável.

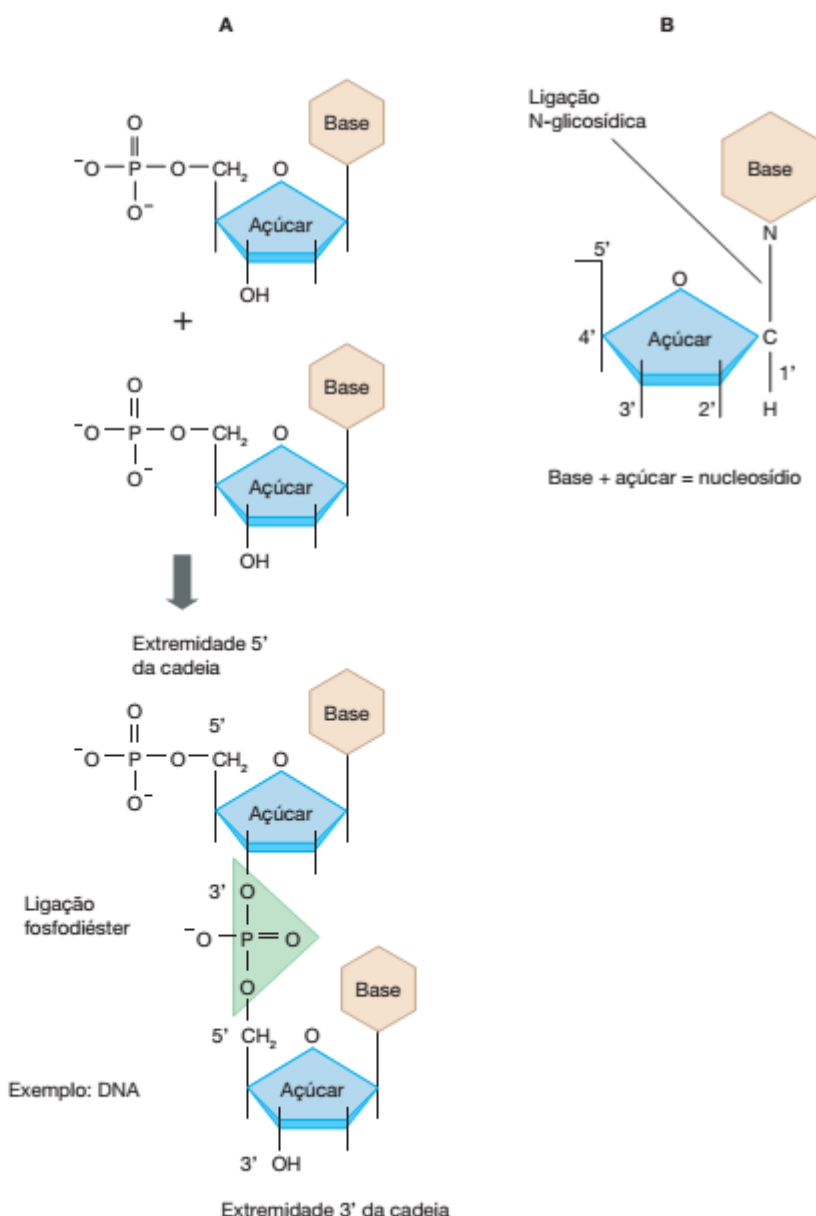



Figura 1.3 Formação da cadeia polinucleotídica. (A) A união dos nucleotídios ocorre por meio de ligações fosfodiéster entre os átomos de carbono 5′ e 3′ da pentose. (B) A base nitrogenada liga-se ao carbono 1′ da pentose, por meio de uma ligação N-glicosídica, formando o nucleosídio.

A maneira como os nucleotídios ligam-se uns aos outros confere às fitas do DNA ou do RNA uma *polaridade*. Assim, em uma extremidade da molécula (*extremidade 5'*), teremos um grupo fosfato livre, e na outra (*extremidade 3'*), um grupo hidroxila livre. A Figura 1.4 ilustra a união de nucleotídios formando um polímero de DNA. Por convenção, ao descrevermos a sequência de bases no DNA ou no RNA, utilizamos o sentido 5′→ 3′.

Ponte de hidrogênio
Ligação química fraca que se forma entre as bases nitrogenadas que pareiam entre si (adenina com timina e citosina com guanina) e mantém as duas fitas de DNA unidas

Dupla hélice
Estrutura química do DNA (duas fitas pareadas e enroladas formando uma hélice)

Em geral, as moléculas de RNA são mais curtas que as de DNA e são estruturas *unifilamentares*, embora, eventualmente, possam dobrar-se na forma de uma alça, em função do pareamento complementar de bases na fita única, estabilizado por pontes de hidrogênio. O DNA, por sua vez, é um filamento duplo constituído por duas cadeias de nucleotídios, mantidas unidas por pontes de hidrogênio que se formam entre bases nitrogenadas complementares. A estrutura tridimensional da molécula de DNA, conforme o modelo proposto por Watson e Crick, consiste em uma dupla hélice helicoidal enrolada ao longo de um eixo comum, com giro para a direita, na qual ambas as fitas são complementares e têm direções opostas, ou seja, são *antiparalelas*. Deste modo, uma das cadeias do DNA tem o sentido 5′→ 3′, e a outra o sentido 3′→ 5′, e a relação espacial entre as duas fitas cria sulcos maiores e menores ao longo de sua extensão. Interações do DNA com proteínas que regulam a expressão gênica evidenciam que algumas delas ligam-se ao sulco maior e outras, ao sulco menor. Os resíduos de fosfato e açúcar, de caráter hidrofílico, ficam situados no lado externo da hélice de DNA, e as bases nitrogenadas, de natureza hidrofóbica, na parte interna da molécula (Figura 1.5A).

O pareamento entre as bases de cada fita de DNA ocorre de maneira padronizada, e as ligações de hidrogênio que se formam entre elas contribuem para manter a dupla hélice unida, bem como conferir estabilidade à molécula. As purinas pareiam com as pirimidinas (adenina

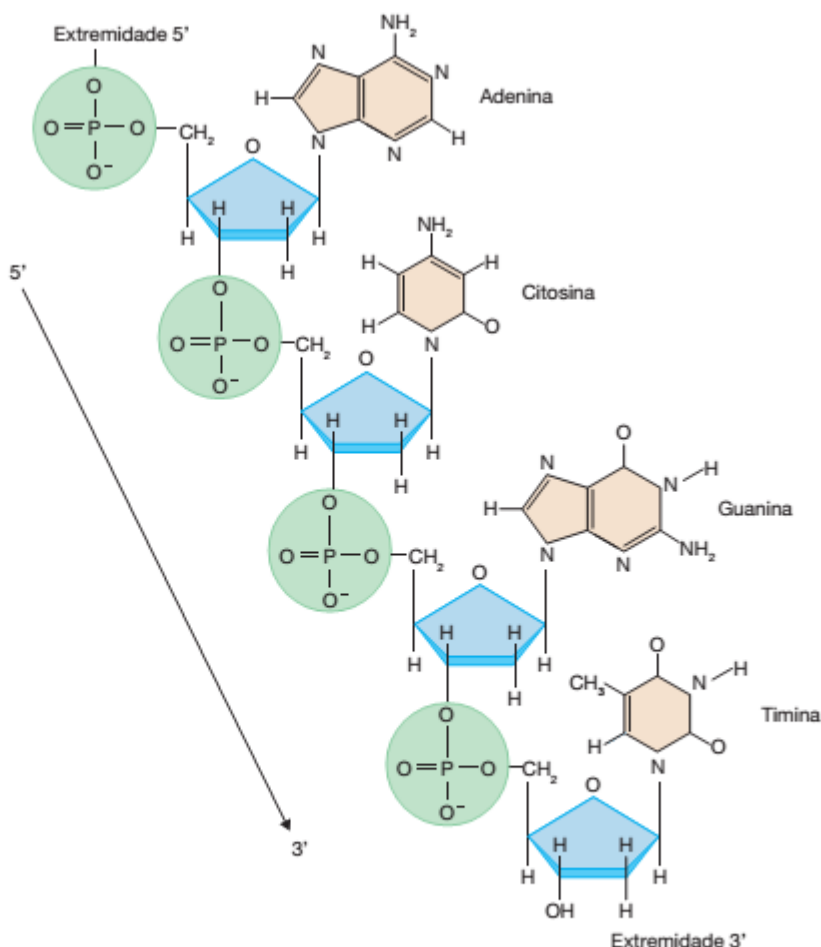



Figura 1.4 Segmento de DNA com quatro nucleotídios. Observe que a molécula tem duas extremidades (5′ e 3′), o que lhe confere uma polaridade.

com timina e citosina com guanina) por meio de pontes de hidrogênio, que são ligações químicas mais fracas se comparadas às ligações açúcar-fosfato. Duas pontes de hidrogênio ligam as bases adenina e timina, e três pontes ligam as citosinas e guaninas (Figura 1.5B).

A especificidade de pareamento existente entre as bases nitrogenadas é a característica estrutural mais importante do DNA e, em função dela, a sequência de bases em uma fita pode ser determinada conhecendo-se a sequência de bases da fita complementar, de modo que a quantidade de adeninas será igual à de timinas e a de citosinas igual à de guaninas.

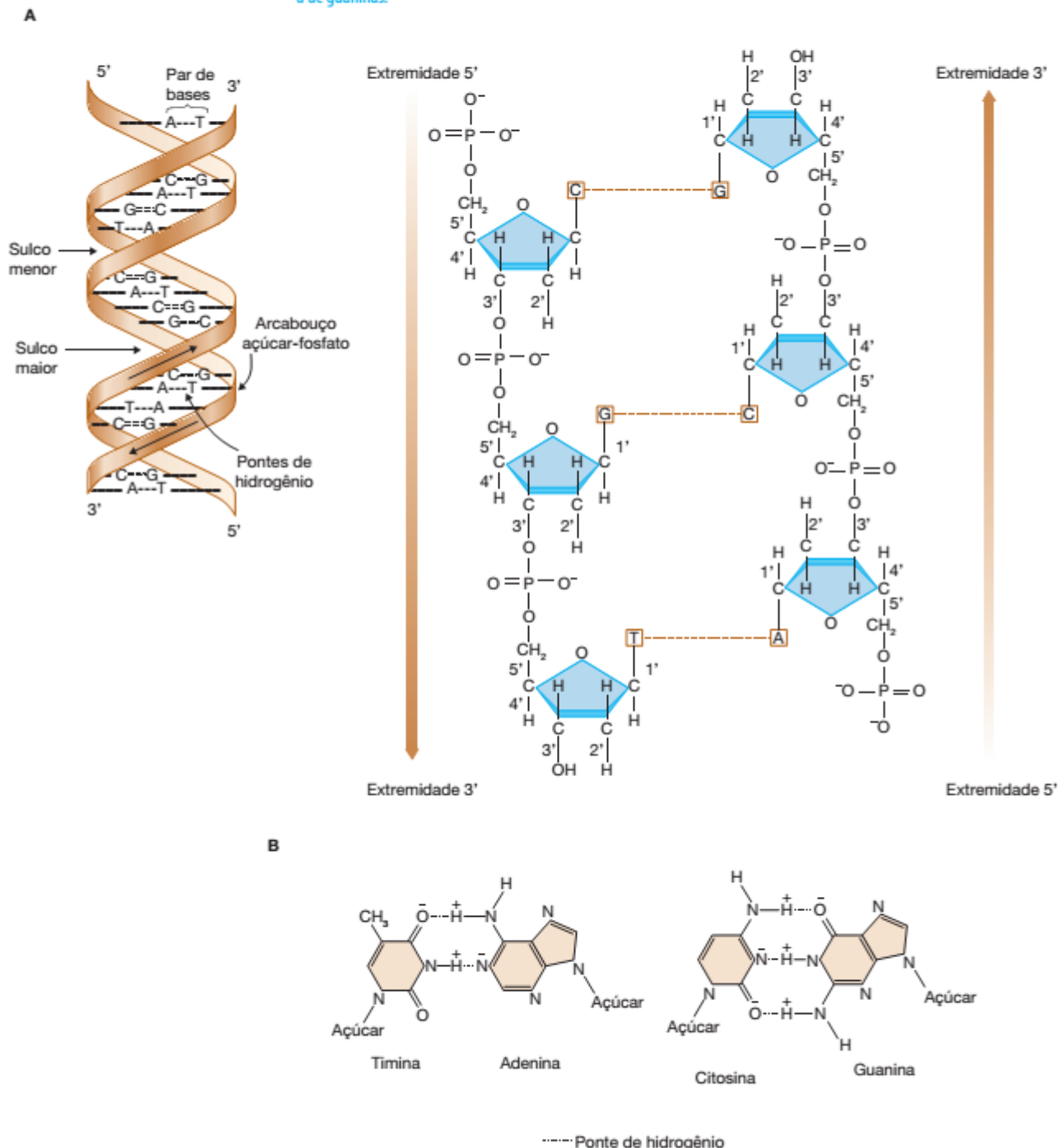



Figura 1.5 Dupla hélice de DNA. (A) Formada por duas fitas complementares de DNA, de sentidos opostos, enroladas na forma de uma hélice, apresentando sulcos maiores e menores. (B) O pareamento entre as bases nitrogenadas, por meio de pontes de hidrogênio, mantém as duas cadeias de DNA unidas.

Formas alternativas da dupla hélice de DNA

Ao contrário de que se imagina, o DNA não é uma molécula invariavelmente estática. Na verdade, a dupla hélice de DNA é uma estrutura dinâmica e flexível, que pode assumir diferentes conformações (*formas A, B* e *Z*), em decorrência de sua interação com outras moléculas (Figura 1.6). A conformação adotada por um determinado segmento da molécula de DNA está intimamente relacionada com a regulação da expressão gênica naquela região. O DNA, como descrito por Watson e Crick, equivale à conformação normal da molécula em condições fisiológicas, conhecida como *DNA B*. A forma B, assim como o *DNA A*, é uma dupla hélice com giro para a direita (dextrógira), porém o DNA B tem 10 pares de bases (pb) a cada volta completa da hélice e é uma molécula mais longa e mais fina (1,9 nm de diâmetro), ao passo que o DNA A, com 11 pares de nucleotídios por volta, é uma dupla hélice mais curta e mais espessa (2,3 nm de diâmetro). O *DNA Z* é uma forma de dupla hélice totalmente distinta das demais: tem rotação à esquerda (levógira), apresenta 12 pb por volta, um único sulco e o esqueleto açúcar-fosfato organiza-se em zigue-zague. O DNA Z é uma hélice com 1,8 nm de diâmetro, e constitui a forma mais longa e mais fina da molécula.

Dextrógira
Refere-se ao giro à direita da dupla hélice de DNA

Levógira
Refere-se ao giro à esquerda da dupla hélice de DNA

Genomas de procariotos e eucariotos

O termo genoma refere-se ao conteúdo total de DNA presente na célula de um determinado organismo, e grandes diferenças na organização genômica de procariotos e eucariotos são observadas. Genomas de mais de 600 organismos já foram sequenciados e muitos outros se encontram em processo de sequenciamento. Em termos de tamanho, eles são extremamente variáveis. Genomas bacterianos têm entre 1 e 6 Megabases (Mb), e, mesmo entre linhagens de uma mesma espécie, como a bactéria *Escherichia coli*, observam-se variações expressivas. A linhagem K12 de *E. coli* tem genoma constituído por pouco mais de 4,6 Mb, ao passo que a linhagem 0157 dessa espécie tem cerca de 5,5 Mb. Entre eucariotos, o genoma da levedura contém, aproximadamente, 13 Mb; o da mosca das frutas, 180 Mb; o do camundongo, 2.600 Mb; e o de humanos, cerca de 3.000 Mb. As plantas evidenciam uma variação extrema, com genomas atingindo até 120.000 Mb.

Genoma
Informação genética total presente na célula de um organismo

Megabase (Mb)
Unidade de comprimento. Assim, 1 Mb equivale a 1 milhão de nucleotídios

Comparando-se a *densidade gênica* dos genomas já sequenciados de alguns organismos, verificamos que eucariotos apresentam menor quantidade de genes por Mb em relação aos procariotos. A menor densidade gênica é observada nos genomas de mamíferos, incluindo o homem. Nas bactérias, a densidade gênica é de, aproximadamente, um gene por quilobase (kb) de DNA, com distância média intergênica de 118 pb, o que contrasta bastante com as baixas densidades gênicas observadas nos genomas de eucariotos, principalmente dos vertebrados. Na verdade, a maior parte do genoma bacteriano é formada por genes que codificam proteínas ou RNAs, e menos de 1% é constituído por elementos repetitivos não codificantes, ao passo

Quilobase (kb)
Unidade de comprimento. Assim, 1 kb equivale a 1.000 nucleotídios

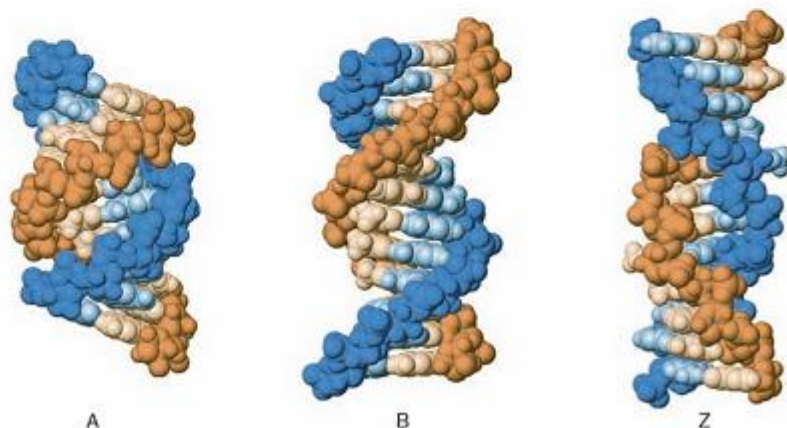



Figura 1.6 Diferentes conformações assumidas pela dupla hélice de DNA (formas A, B e Z).

DNA repetitivo
Sequência de DNA presente em múltiplas cópias no genoma

que em eucariotos a maior parte do DNA não codifica proteínas ou moléculas maduras de RNA. A baixa densidade gênica observada nos genomas de eucariotos mais complexos está diretamente correlacionada com a grande quantidade de DNA repetitivo presente nestes organismos, como veremos a seguir.

Em procariotos, o genoma é formado em sua quase totalidade por sequências únicas de DNA, que, por definição, são aquelas que ocorrem em pequeno número na célula (1 a 10 cópias). Diferentemente, a organização genômica em eucariotos é bem mais complexa, e nestes podemos distinguir dois outros tipos de sequências de DNA, além das sequências de cópia única. Uma quantidade expressiva do genoma eucariótico é formada por DNA repetitivo, representado por sequências moderadamente repetitivas (10 a 10^5 cópias), presentes em maior quantidade, e por sequências altamente repetitivas (mais de 10^5 cópias). A disposição dessas sequências repetidas no genoma é variável. Elas podem estar dispostas lado a lado (*em tandem*) ou intercaladas com sequências de copia única. Análises genômicas revelam que cerca de 45% dos genomas de mamíferos são constituídos por sequências de DNA moderadamente repetitivas, derivadas de elementos de transposição. Um elemento transponível é um segmento de DNA que tem a capacidade, durante o processo de transposição, de se deslocar de um local para outro do genoma.

Elemento de transposição (elemento genético transponível)
Sequência de DNA, presente em procariotos e eucariotos, capaz de se mover de um local para outro no genoma

O genoma humano | DNA nuclear e mitocondrial

A informação genética total existente na célula humana engloba o DNA nuclear (*nDNA*) e o DNA das mitocôndrias presentes no citoplasma (*mtDNA*), de modo que uma perfeita interação entre estes dois genomas se faz necessária para que a célula realize suas funções adequadamente.

Diploide
O que tem duas cópias de cada cromossomo. As células somáticas humanas são diploides, ou seja, 2n

O genoma nuclear humano

O *genoma nuclear humano*, com cerca de 6 bilhões de nucleotídios por núcleo diploide, representa mais de 99,95% de todo o DNA presente em uma célula humana e tem uma estrutura bem mais complexa em relação ao genoma mitocondrial. O *nDNA* encontra-se distribuído em 46 cromossomos lineares (22 pares de autossomos e 2 c*romossomos sexuais X e Y*), e cada um deles é constituído por uma única molécula de DNA. Em função de cada cromossomo humano, de uma extremidade à outra, ser formado por uma dupla hélice contínua de DNA, ele é denominado unifilamentar ou uninêmico. A dupla fita de DNA, por sua vez, está intimamente associada a proteínas específicas, formando a cromatina, complexo de DNA e proteínas que forma os cromossomos eucarióticos. Os genes ficam dispostos ao longo dos cromossomos em locais definidos, denominados *loci* (no singular, *locus*).

Cromossomo
Estrutura que contém a informação genética. Os cromossomos de eucariotos são estruturas lineares formadas por DNA e proteínas associadas. O cromossomo bacteriano é constituído por uma molécula de DNA circular

Autossomo
Qualquer cromossomo nuclear diferente dos cromossomos sexuais X e Y. Existem 22 pares de autossomos nas células somáticas humanas

Cerca de 50% do genoma humano é constituído por *sequências de DNA de cópia única*, que contêm a maioria dos genes codificadores de proteínas. A outra metade de nosso genoma é formada por diferentes classes de *DNA repetitivo*, não codificante e de função desconhecida em sua quase totalidade. As sequências repetidas no DNA podem apresentar-se agrupadas (em tandem) ou espalhadas pelo genoma, intercaladas por sequências de cópia única. Em função do comprimento da unidade de repetição, de sua localização e organização no genoma e do tamanho total da série de unidades repetidas, as seguintes categorias de DNA repetitivo são distinguidas:

1. *DNA altamente repetitivo*. Constitui cerca de 5% do genoma humano. Esta classe de sequências repetidas é representada pela família do *DNA satélite*, presente em regiões cromossômicas específicas adjacentes aos centrômeros (regiões centroméricas). Em humanos, o DNA satélite é constituído por uma sequência de 171 pb, repetida em tandem numerosas vezes na região centromérica de cada cromossomo, denominada *DNA satélite alfa* ou *alfoide*. O *DNA satélite alfa* é crucial para que a segregação cromossômica ocorra de maneira precisa durante a divisão celular.
2. *DNA moderadamente repetitivo*. Representa a maior parte do DNA repetitivo. Nesta classe, incluem-se diferentes tipos de repetições que se encontram dispersas pelo genoma humano:

Centrômero
Constrição primária existente nos cromossomos eucarióticos que mantém as cromátides irmãs unidas e delimita os braços cromossômicos (braço curto e braço longo)

DNA satélite
Fração do DNA genômico formada por sequências altamente repetidas que pode ser separada do DNA total por centrifugação

longas sequências repetidas em tandem (minissatélites), curtas sequências repetidas em tandem (microssatélites) e *elementos de transposição*. Além desses tipos de sequências, uma pequena fração do DNA humano moderadamente repetitivo inclui alguns genes funcionais que, diferentemente da maioria dos genes codificadores representados por sequências de DNA de cópia única, encontram-se presentes no genoma em múltiplas cópias em tandem. Como exemplo, temos os genes codificantes de rRNA, moléculas necessárias em grandes quantidades na célula. Cabe salientar, entretanto, que em sua maioria o DNA repetitivo não tem função codificadora.

O DNA minissatélite, exemplificado pelos VNTR (número variável de repetições em tandem), é formado por sequências de 10 a 100 nucleotídios que se repetem lado a lado inúmeras vezes, formando blocos de repetições com extensão entre 1 e 20 kb, com extrema variação entre indivíduos. Os *VNTR* encontram-se espalhados no genoma, presentes dentro de genes e em regiões intergênicas, e, por serem altamente polimórficos, são amplamente usados como marcadores moleculares em análises genômicas.

Outra fração do DNA moderadamente repetitivo é representada pelos microssatélites, formados por unidades com pequeno número de nucleotídios (< 10), em sua maioria *dinucleotídios*, *trinucleotídios* ou *tetranucleotídios*, que se repetem em tandem (STR). De modo semelhante aos *VNTR*, os *STR* também são marcadores moleculares altamente polimórficos, com ampla variação interindividual, extensamente utilizados para fins diversos.

Inúmeras repetições em tandem de uma curta sequência de bases (5′-TTAGGG-3′, em humanos e em outros vertebrados) também estão presentes nas extremidades dos cromossomos eucarióticos, formando uma estrutura denominada telômero, que protege as pontas cromossômicas da degradação enzimática por endonucleases e da fusão com outros cromossomos. Quebras nas extremidades cromossômicas tornam estas coesivas, dando origem, geralmente, a cromossomos dicêntricos ou cromossomos em anel. A replicação dos telômeros, como veremos adiante, é uma etapa à parte na duplicação da molécula de DNA como um todo, pois requer a participação de uma enzima especial, a telomerase. Na ausência dessa enzima, o DNA repetitivo telomérico torna-se cada vez mais curto a cada ciclo de replicação.

Os telômeros são essenciais na estabilidade e na integridade cromossômica, já que impedem a união entre as pontas dos cromossomos eucarióticos.

Além do DNA *minissatélite* e *microssatélite*, a maior parte da fração das sequências moderadamente repetitivas presentes no genoma humano é representada por *sequências de DNA móveis* ou *elementos de transposição*, que são capazes de se inserir em diferentes locais no genoma. Dois tipos principais de retrotransposons, os elementos nucleares intercalados curtos (*SINE*), com menos de 500 pb de extensão, e os elementos nucleares intercalados longos (*LINE*), com 6 kb de comprimento, estão presentes no genoma humano. Entre os SINE, a família de *elementos Alu* é o tipo de repetição mais abundante, e entre os LINE, a família *L1* é a sequência móvel mais importante no genoma humano. A presença de *elementos de transposição* em uma fração significativa dos genomas de eucariotos reforça a ideia de que eles desempenharam um importante papel ao longo do processo evolutivo.

Uma quantidade substancial do genoma nuclear humano é de natureza repetitiva e essas sequências de DNA não têm função codificante, exercendo um papel ainda desconhecido.

Os dados revelam que apenas uma pequena fração do nDNA humano, menos de 2%, codifica proteínas. Considerando que o número de proteínas sintetizadas pelo ser humano é bem maior do que o de genes que as codificam, como explicar a grande diversidade proteica face ao número reduzido de genes humanos que são transcritos em mRNA e traduzidos em polipeptídios?

Na verdade, muitos genes nucleares humanos são capazes de originar não apenas uma, mas várias proteínas.

Os genes contidos no nDNA, em sua maioria, são formados por *regiões codificantes* (éxons), com tamanho médio de cerca de 200 pb, intercaladas por *regiões não codificantes* (íntrons), que exibem enorme variação de comprimento. Cabe ressaltar que, enquanto o gene da proteína muscular *conectina* (ou *titina*) tem mais de 300 íntrons, outros genes nucleares de pequeno

Telômero
Estrutura que protege as extremidades dos cromossomos eucarióticos lineares e contribui para a integridade e a estabilidade cromossômica. É formado por curtas sequências de DNA repetidas em tandem inúmeras vezes

Endonuclease
Grupo heterogêneo de enzimas que clivam ligações entre nucleotídios nas fitas de DNA ou de RNA

Cromossomo dicêntrico
Cromossomo com dois centrômeros

Cromossomo em anel
Cromossomo estruturalmente anormal, no qual os telômeros de ambos os braços foram deletados e as duas extremidades se uniram, originando uma conformação circular, ou seja, em anel

Telomerase
Enzima que adiciona sequências de DNA telomérico às extremidades dos cromossomos lineares de eucariotos

Retrotransposon
Sequência de DNA móvel capaz de se transpor para diferentes locais no genoma

Éxon
Região codificante de um gene eucariótico, que, após transcrição e tradução, estará representada na cadeia polipeptídica codificada pelo gene

Íntron
Região gênica não codificante, que se intercala às regiões codificadoras e que é removida durante o *splicing* do mRNA, não estando representada no polipeptídio codificado pelo gene.

Gene
Unidade física e funcional da hereditariedade, formada por um segmento de DNA específico (ou de RNA, em alguns vírus)

RNA mensageiro (mRNA)
Molécula de RNA transcrita do DNA que será traduzida na sequência de aminoácidos de um polipeptídio

Polipeptídio
Cadeia de aminoácidos unidos por ligações peptídicas

RNA ribossômico (rRNA)
Molécula de RNA que compõe as subunidades ribossômicas. O ribossomo é uma estrutura formada por duas subunidades de ribonucleoproteínas, na qual a síntese de proteínas ocorre

RNA transportador (tRNA)
Molécula de RNA que participa do processo de tradução do mRNA, transportando aminoácidos específicos durante a síntese de proteínas

Pequeno RNA nuclear (snRNA)
Pequenas moléculas de RNA presentes no núcleo das células eucarióticas que se ligam a proteínas e atuam no processamento do pré-mRNA

Pequeno RNA nucleolar (snoRNA)
Pequenas moléculas de RNA que participam da maturação do rRNA

Encadeamento (*splicing*)
Parte do processo de maturação do pré-mRNA. Refere-se à retirada dos íntrons e união dos éxons, originando uma molécula de mRNA constituída somente por regiões codificantes

Região subtelomérica
Região cromossômica proximal ao telômero, com elevada concentração gênica

G·F Genética em Foco

O genoma nuclear humano

O sequenciamento do genoma nuclear humano não foi uma tarefa fácil, face a sua complexa arquitetura, com uma grande quantidade de diferentes tipos de sequências de DNA que se repetem. Na verdade, foi exigido um intenso esforço conjunto da comunidade científica mundial. No ano de 2001, pouco mais de uma década após o início do *projeto genoma humano*, foram divulgados os primeiros rascunhos sobre o número de genes codificantes em nosso genoma, indicando números bem menores em relação àqueles inicialmente supostos (entre 80.000 e 100.000 genes). Em consulta aos principais catálogos de genes humanos, observamos que as estimativas vêm passando por aferições sucessivas, e que as mais atuais indicam a existência de cerca de *30.000 genes funcionais*. Destes, 21.701 codificam proteínas, ou seja, são transcritos em RNA mensageiro (mRNA) e traduzidos em polipeptídios, e 8.483 são transcritos em moléculas maduras de RNA, não codificantes, de funções diversas. Alguns RNAs auxiliam no processo de expressão gênica, como o RNA ribossômico (rRNA) e o RNA transportador (tRNA), que participam da síntese de proteínas, como veremos em detalhes no Capítulo 3. Além destes, temos os pequenos RNA nucleares (snRNA) e os pequenos RNA nucleolares (snoRNA). Os *snRNA* associam-se a proteínas específicas e atuam no processo de encadeamento (*splicing*) do mRNA, e os *snoRNA* participam do processo de maturação do rRNA.

tamanho, como os codificadores das proteínas *histonas*, não têm íntrons. Cerca de 60% dos genes nucleares codificam moléculas de mRNA que podem ser processadas de diferentes modos, por meio da recomposição alternativa de éxons e íntrons, gerando um número de proteínas bem maior do que o número de genes existentes. No Capítulo 3, veremos em detalhes como o processo de encadeamento do mRNA ocorre.

Outra característica dos genes humanos é que eles são extremamente variáveis em tamanho. Alguns, como os que codificam a insulina e a β-globina, têm menos de 2 kb de comprimento; outros apresentam entre 10 e 100 kb de tamanho, como o gene do *colágeno tipo VII*, com 31 kb, e o da enzima *fenilalanina hidroxilase*, com 90 kb. Existem, ainda, aqueles com tamanho superior a 100 kb. Nessa classe, incluem-se o gene do *fator VIII*, com 186 kb, o gene *CFTR* (associado à fibrose cística), com 250 kb, e o gene da *distrofina*, relacionado com as distrofias musculares de Duchenne e de Becker, que é o maior gene humano conhecido, com pouco mais de 2.200 kb de comprimento.

Observa-se também que a distribuição gênica não ocorre de maneira uniforme nos 22 autossomos e nos cromossomos sexuais humanos. Embora existam, em média, 6,6 genes a cada Mb do genoma humano, a densidade gênica varia substancialmente entre cromossomos, e mesmo entre regiões cromossômicas. Particularmente, as regiões subteloméricas apresentam uma alta densidade gênica. Alguns cromossomos, como o 19 e o 17, são ricos em genes e têm, respectivamente, cerca de 22 e 15 genes por Mb. Outros, como os cromossomos 4, 13 e 18, têm em torno de 4 genes por Mb, ou seja, uma baixa densidade gênica. Um caso particular é o do cromossomo Y, que tem a menor densidade gênica, 1,4 gene por Mb.

Análises comparativas de genomas revelam que mais de 50% dos genes humanos compartilham significante homologia com genes de outras espécies, como o cachorro (*Canis familiaris*), o macaco *rhesus* (*Macaca mulatta*) e o chimpanzé (*Pan troglodytes*). Observa-se maior similaridade com o chimpanzé (cerca de 98%). Quando genes de diferentes espécies apresentam sequências de bases semelhantes, eles são denominados genes ortólogos.

O genoma mitocondrial humano

As *mitocôndrias* são organelas citoplasmáticas existentes na quase totalidade das células de eucariotos, e sua função principal é fornecer a energia necessária ao metabolismo celular. Cada célula humana contém múltiplas mitocôndrias, e estas, por sua vez, têm várias cópias de *mtDNA*. O número de cromossomos mitocondriais é maior nas células que demandam altos gastos energéticos, como as nervosas e musculares.

Ovócito
Célula germinativa feminina haploide ou gameta feminino

Zigoto
Célula diploide formada pela união de dois gametas haploides (ovócito e espermatozoide) a partir da qual o embrião se desenvolve

Espermatozoide
Célula germinativa masculina haploide ou gameta masculino

Diferentemente do genoma nuclear, o *genoma mitocondrial humano* é extremamente compacto e 93% de seu DNA apresenta função codificadora. O cromossomo mitocondrial é formado por uma pequena molécula de DNA circular dupla fita, de 16.569 pb de comprimento, composta por uma cadeia rica em bases G (fita pesada ou H) e uma rica em bases C (fita leve ou L). O genoma presente nas mitocôndrias de todo ser humano, homem ou mulher, é definido exclusivamente pelo genoma do ovócito, pois, na formação do zigoto, o espermatozoide contribui apenas com o genoma nuclear e não com o mitocondrial. Por essa razão, dizemos que o *mtDNA* é herdado pela linhagem materna (herança matrilinear). O cromossomo mitocondrial contém apenas *37 genes*, dos quais 13 codificam 15% das proteínas que participam da cadeia respiratória, sendo a maioria das proteínas com funções específicas nas mitocôndrias codificadas pelo DNA nuclear. Os demais 24 genes mitocondriais codificam moléculas de RNA (22 codificam tRNA e 2 codificam rRNA). Nenhum dos 37 genes mitocondriais contém íntrons e alguns deles apresentam superposição de sequências.

Organização do DNA nos cromossomos

Poliploide
O que tem mais de duas cópias de cada cromossomo, ou seja, triploide (3n), tetraploide (4n), …

Em contraste com as bactérias, nas quais a informação genética está contida em uma única dupla fita de DNA, formando um cromossomo circular, na maioria dos eucariotos o DNA encontra-se distribuído em diferentes estruturas lineares (*cromossomos*), presentes em duas cópias, como nos seres humanos (*diploides*), ou em várias cópias, como em algumas plantas que florescem (poliploides). Contrariamente ao DNA bacteriano, que está associado a uma pequena quantidade de proteínas, nas células eucarióticas as moléculas de DNA que formam os cromossomos estão conjugadas a uma quantidade expressiva de proteínas, constituindo a *cromatina*. A estrutura da cromatina é bastante dinâmica. Diferentes níveis de organização de sua estrutura são notados, de modo que a cromatina pode apresentar-se na forma compactada ou relaxada. A reversibilidade observada na estrutura da cromatina está diretamente relacionada com o controle da expressão gênica.

Histona
Grupo de proteínas básicas, evolutivamente conservadas, que se ligam ao DNA formando a cromatina

Compactação do DNA nos cromossomos de eucariotos

Sabendo que o núcleo de uma célula humana mede, aproximadamente, 6 μm de diâmetro, como é possível ele acomodar todo o genoma nuclear humano, considerando que as moléculas de DNA que formam os 46 cromossomos, quando alinhadas, medem cerca de 2 metros de extensão? Isso ocorre em função da capacidade de compactação do DNA.

Meiose
Tipo de divisão celular que resulta na formação dos gametas, células haploides (n) com metade do número cromossômico da célula-mãe (2n)

Mitose
Tipo de divisão celular que resulta na produção de duas células filhas geneticamente iguais à célula-mãe

Interfase
Período entre duas mitoses sucessivas

O empacotamento do DNA no núcleo das células humanas e de eucariotos em geral ocorre em diferentes níveis, especificamente por meio da ligação dessa molécula com proteínas básicas específicas, denominadas histonas, cuja massa total na cromatina é equivalente à massa de DNA. A associação do DNA às *histonas* possibilita que a cromatina se apresente distendida, bastante compactada ou parcialmente condensada. Em linhas gerais, a cromatina mostra-se mais extensamente condensada durante a divisão celular (meiose ou mitose), quando os cromossomos atingem níveis de compactação da ordem de 10^4 vezes, do que no período de não divisão (interfase), quando não é possível distinguir os cromossomos em função de a cromatina estar muito distendida.

A Figura 1.7 mostra os diferentes níveis de condensação da cromatina, desde a sua organização nos *nucleossomos* até a formação do cromossomo mitótico. A unidade estrutural básica da cromatina eucariótica é o nucleossomo, que consiste em um núcleo proteico formado por oito moléculas de histonas (duas de cada uma das proteínas histonas H2A, H2B, H3 e H4), em torno do qual a dupla fita de DNA está enrolada. Precisamente, 147 nucleotídios do DNA dão 1,7 vol-

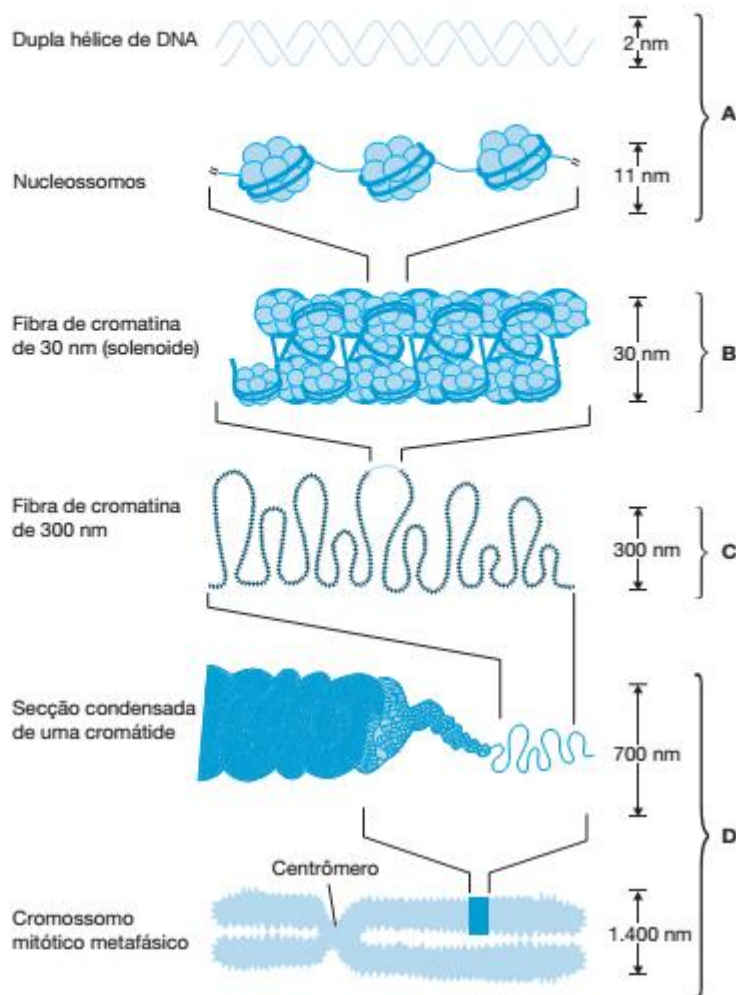



Figura 1.7 Níveis de compactação da cromatina. (A) Estrutura dos nucleossomos. (B) Fibra de 30 nm ou solenoide. (C) Fibra de 300 nm. (D) Cromátide do cromossomo metafásico.

ta ao redor desse *octâmero de histonas* (Figura 1.7A). Estas são proteínas cujas sequências de aminoácidos apresentam-se evolutivamente conservadas e que, em função de terem um grande percentual de aminoácidos de carga positiva, são capazes de ligar-se firmemente à dupla hélice do DNA. Entretanto, essa associação DNA-histonas é reversível. Na verdade, diferentes mecanismos, como a ligação da cromatina a proteínas não histonas ou modificações químicas das histonas, são capazes de reduzir, temporariamente, a afinidade das histonas pelo DNA, alterando a organização da cromatina, o que possibilita o acesso ao DNA de inúmeros fatores reguladores da expressão gênica. Deste modo, genes presentes em um segmento de DNA não associado às histonas podem se expressar, ou seja, ser transcritos em moléculas de RNA. Diferentemente das histonas, as proteínas não histonas são muito heterogêneas, com ampla variação entre as espécies e mesmo entre diferentes tipos celulares de um mesmo organismo.

Cada nucleossomo é separado do outro por um segmento de DNA espaçador que os une, formando uma fibra de 11 nm de diâmetro que, ao ser visualizada por microscopia eletrônica, se assemelha às contas de um colar (Figura 1.7A). Uma célula humana diploide contém cerca de 30 milhões de *nucleossomos*, e eles representam o *primeiro nível de compactação da cromatina*. A histona H1 liga-se à fibra de 11 nm, na região de espaçamento entre os nucleossomos, e promove seu dobramento, ocasionando o empilhamento dos nucleossomos e a formação da fibra de 30 nm, denominada solenoide (Figura 1.7B). Esta, por sua vez, pode formar alças secundárias, aumentando o nível de empacotamento da cromatina, que alcança 300 nm de diâmetro (Figura 1.7C) e atinge seu grau máximo de condensação na metáfase da mitose, quando cada cromátide cromossômica tem uma espessura de 700 nm (Figura 1.7D).

Nucleossomo
Unidade estrutural básica da cromatina eucariótica que consiste em um segmento de DNA enrolado ao redor de um octâmero de histonas e representa o primeiro nível de compactação da cromatina

Solenoide
Fibra de 30 nm de diâmetro formada pelo empilhamento dos nucleossomos

Cromátide
Filamento de cromatina que se forma pela autoduplicação do DNA. Duas cromátides unidas pelo centrômero formam o cromossomo

Heterocromatina
Parte da cromatina mais condensada, intensamente corada durante a interfase.

Eucromatina
Parte da cromatina mais distendida, levemente corada durante a interfase.

Em função do grau de compactação exibido pela cromatina, esta pode ser denominada heterocromatina, representando a cromatina altamente condensada e transcricionalmente inativa, ou eucromatina, em sua forma menos compactada. A eucromatina é a parte da cromatina mais distendida que se mostra levemente corada durante a interfase. A maioria dos genes estruturais localiza-se nas regiões eucromáticas. Na eucromatina os genes serão transcritos em função da acetilação das proteínas histonas e da fraca ligação da histona H1, que permitirão o acesso ao DNA da maquinaria proteica e enzimática necessária para a transcrição gênica.

Transmissão da informação genética

Ciclo celular

Célula somática
Refere-se a qualquer célula de um organismo diferente das germinativas ou dos gametas. As células que formam os tecidos do corpo são somáticas

O conhecimento de que *uma célula se origina a partir de outra célula* nos leva a concluir que a continuidade da vida está fundamentada na propagação das células. O crescimento e a reprodução de organismos multicelulares têm como premissa a divisão celular, por meio da qual uma célula somática, ao se dividir, origina duas células geneticamente iguais. Para que isso ocorra, é necessário que o DNA da célula-mãe seja previamente replicado e, então, repartido entre as novas células formadas.

Ciclo celular
Evento celular cíclico que envolve a divisão celular e a interfase (período entre duas divisões)

A replicação do DNA nas bactérias, quando estas estão em condições ambientais favoráveis, ocorre de modo contínuo. Em menos de 60 min, a bactéria *E. coli* é capaz de realizar todo o seu ciclo de vida, ou seja, duplicar seu único cromossomo com 4,6 Mb e se dividir em duas células. Assim, em algumas horas, ela pode originar uma colônia com milhares de células. Diferentemente das bactérias, em eucariotos o processo de divisão celular é mais complexo e consome mais tempo, em parte porque o conteúdo de DNA a ser replicado é maior e também em função de ele estar organizado na forma de cromatina.

Interfase
Período do ciclo celular entre duas divisões celulares

Citocinese
Divisão do citoplasma que resulta nas células-filhas durante a mitose ou a meiose

Todas as células eucarióticas passam por diferentes fases, que, em conjunto, constituem o ciclo celular e garantem a transmissão da informação genética de uma célula para outra. Em linhas gerais, o ciclo celular compreende dois estágios distintos que se sucedem: um período mais longo de não divisão (interfase), composto pelas fases G_1 (G = *gap*, intervalo; primeiro intervalo), S (síntese do DNA) e G_2 (segundo intervalo), e a divisão celular propriamente dita (fase M), que inclui a *mitose* e a citocinese (Figura 1.8). As fases G_1 e G_2 são intervalos que ocorrem entre a duplicação do DNA e a divisão celular.

A célula permanece a maior parte de sua vida em interfase, que representa cerca de 90% do ciclo celular. Durante essa fase ocorre intensa atividade metabólica, crescimento e diferenciação celular e replicação do DNA cromossômico, preparando a célula para entrar em divisão.

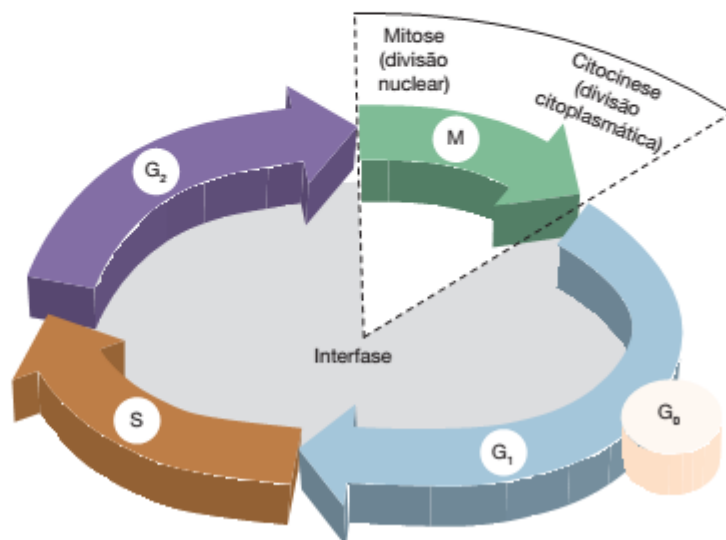



Figura 1.8 Fases do ciclo celular em eucariotos. Interfase ($G_1/G_0 \rightarrow S \rightarrow G_2$) e fase M (mitose e citocinese). G_1: crescimento e metabolismo; S: replicação do DNA; G_2: preparação para a mitose; M: distribuição igualitária dos cromossomos duplicados para as células-filhas (mitose) e divisão celular propriamente dita, originando duas células filhas (citocinese).

O período G_1 representa o estado normal da célula, constituindo um período de crescimento e metabolismo celular intensos. Essa fase está intimamente relacionada com o controle da proliferação celular, pois nessa etapa algumas células podem interromper o ciclo e entrar em repouso (G_0) ou permanecer em G_1, preparando a célula para a replicação do DNA que ocorrerá na fase seguinte (S). Células em repouso (G_0) apenas não proliferam, embora se mantenham metabolicamente ativas. Algumas células permanecem definitivamente em repouso e outras são estimuladas para dar continuidade ao ciclo, ou seja, retomarem a fase G_1. Cabe observar que em G_1 cada cromossomo é constituído por uma única dupla hélice de DNA, e, após a duplicação dessa molécula, no final da fase S, cada cromossomo terá duas cromátides-irmãs, cada uma delas formada por uma cópia da dupla hélice de DNA original. Desta forma, em G_2, período que antecede à mitose, todo o conteúdo de DNA encontra-se duplicado, bem como a massa celular total. A fase G_2 termina quando a célula inicia o curto estágio de divisão celular (fase M), produzindo duas células-filhas com o mesmo conteúdo de DNA originalmente presente na célula-mãe, com cada uma delas apta a repetir o ciclo. As células eucarióticas, em geral, alternam ciclos de divisão e não divisão.

A duração do ciclo celular, como um todo, e também o período de cada fase podem variar amplamente de acordo com o tipo de célula. Em humanos, existem ciclos de curta duração nas células que se dividem de maneira muito rápida, como as embrionárias, sendo as fases G_1 e G_2 breves ou inexistentes. Em alguns tecidos nos quais ocorre uma contínua reposição de células, como a epiderme e a mucosa intestinal, o ciclo celular dura poucas horas. Por outro lado, em tecidos adultos que se desenvolvem lentamente, o ciclo pode se estender por vários meses. Algumas células, como os neurônios, quanto inteiramente diferenciadas, param de se dividir e ficam em repouso (G_0). Em geral, o ciclo celular de uma célula humana dura, aproximadamente, 24 h. A fase S transcorre em torno de 8 a 9 h, a mitose dura cerca de uma hora e os intervalos entre a duplicação do DNA e a divisão celular são variáveis. A fase G_2 dura de 3 a 4 h e G_1 é a etapa mais variável do ciclo celular, podendo prolongar-se por algumas horas (cerca de 10) em determinadas células e dias ou meses em outras. As células permanecem a maior parte de sua existência na fase G_1 ou na fase G_0, quando em repouso.

A progressão das células eucarióticas pelas diferentes fases do ciclo é um processo de alta complexidade e estritamente regulado, que envolve a participação de proteínas específicas. Danos na atividade dessas proteínas fatalmente levam ao descontrole da divisão celular, fazendo com que as células passem a se dividir de maneira anormal, o que pode ocasionar câncer, como veremos detalhadamente no Capítulo 9.

Duplicação do DNA

A passagem da informação biológica contida no DNA, por meio das gerações, requer que esta seja transmitida de modo preciso a cada divisão celular. Isso se dá pela replicação dessa molécula por um mecanismo que ocorre de maneira coordenada com a divisão celular e envolve a participação de um grande aparato enzimático.

A especificidade do pareamento de bases na dupla hélice do DNA constitui o fundamento de sua duplicação, pois torna possível que cada cadeia sirva de molde para a síntese de uma nova fita complementar.

Embora muito de nosso conhecimento sobre como a replicação do DNA ocorre advenha de estudos em bactérias, as etapas desse processo são basicamente semelhantes em procariotos e eucariotos, incluindo humanos, apesar de mais complexas nestes últimos. Isso decorre, principalmente, de três aspectos: a grande quantidade de DNA presente na célula eucariótica em relação à procariótica; a associação do DNA nuclear de eucariotos a proteínas, formando a cromatina; os cromossomos eucarióticos serem estruturas lineares e não circulares.

O início da replicação: origens, forquilhas e unidades de replicação

Onde se inicia a replicação de uma molécula de DNA? Em uma região aleatória ou em um local específico?

Origem de replicação
Região da molécula de DNA onde a duplicação se inicia

Bolha de replicação
Estrutura semelhante a uma bolha que se forma na região onde o DNA está sendo duplicado, em função da separação das duas fitas da molécula

Unidade de replicação (réplicon)
Segmento de DNA que está sendo duplicado

Forquilha de replicação
Bifurcação formada na extremidade da bolha de replicação, em função do desenrolamento da dupla hélice, na qual o DNA está sendo duplicado

Duplicação semiconservativa
Característica da replicação do DNA. A duplicação desta molécula origina duas novas moléculas, cada uma formada por uma fita de DNA molde e uma fita de DNA recém-sintetizada

Replissomo
Complexo sistema proteico-enzimático que atua de maneira coordenada em cada réplicon permitindo que a dupla hélice se desenrole e se mantenha separada durante o processo de replicação

Proteína SSB
Aquela que se liga e estabiliza as fitas simples de DNA na região da forquilha de replicação para que elas possam servir de molde para a síntese das novas cadeias

DNA helicase
Enzima que desenrola e separa a dupla hélice de DNA na forquilha de replicação

DNA topoisomerase
Enzima que promove quebras transitórias no DNA, auxiliando seu desenrolamento e impedindo a super-helicoidização na região anterior à forquilha de replicação

As evidências mostram claramente que o processo de duplicação do DNA tem início em uma região específica da molécula, chamada origem de replicação, e, nesta, a dupla fita de DNA forma a bolha de replicação. A presença de uma única origem de replicação é característica do cromossomo circular bacteriano, ou seja, nas bactérias a síntese do DNA inicia-se em um único ponto, denominado *oriC*, e progride em ambas as direções, de maneira que no cromossomo bacteriano observamos apenas uma unidade de replicação (réplicon) (Figura 1.9A).

Os cromossomos lineares de eucariotos, constituídos por grandes quantidades de DNA, diferentemente dos procariotos, necessitam ser duplicados em um período de tempo específico do ciclo celular (fase S). Desta forma, ao longo de sua extensão eles apresentam *múltiplas origens de replicação* e cada uma delas controla a síntese do DNA em um réplicon. A existência de várias unidades de replicação possibilita que a síntese das novas fitas de DNA ocorra, de maneira simultânea, a partir de diferentes origens ao longo da molécula, formando inúmeras bolhas de replicação, que progridem bidirecionalmente, garantindo maior rapidez ao processo de duplicação. Se um cromossomo eucariótico típico tivesse uma única origem de replicação, sua duplicação levaria dias para ser concluída. De fato, genomas eucarióticos inteiros são replicados em questão de horas. A levedura do pão (*Saccharomyces cerevisiae*) tem entre 250 e 400 réplicons, e a replicação dos bilhões de pares de bases que formam o genoma nuclear humano envolve a formação de cerca de 10.000 réplicons distribuídos ao longo dos 46 cromossomos. Durante todo o processo de replicação do DNA, tanto em procariotos quanto em eucariotos, a forquilha move-se, a partir da origem de replicação, em ambas as direções ao logo da molécula, ou seja, *a duplicação do DNA é bidirecional*. Assim, nos cromossomos eucarióticos lineares, a partir das múltiplas origens de replicação que se formam, a duplicação prossegue em ambas as direções até os réplicons adjacentes se fundirem e toda a molécula de DNA ser duplicada (Figura 1.9B). Nas extremidades da bolha de replicação, a dupla hélice de DNA se desenrola e forma bifurcações, denominadas forquilhas de replicação.

Em todos os organismos vivos, a replicação do DNA resulta na formação de *dois novos heterodúplices de DNA*, constituídos por uma fita molde e uma fita recém-sintetizada. Por isso, essa molécula tem duplicação semiconservativa (Figura 1.9C).

O início do processo de replicação semiconservativa do DNA, cuja estrutura, como vimos, é uma dupla fita helicoidal, requer, além do reconhecimento das regiões de origem, o desenrolamento da hélice. Para que ocorra a síntese das novas cadeias complementares, é necessário que as duas fitas parentais se separem, formando as *forquilhas de replicação*. Esse processo envolve a participação de um complexo sistema proteico-enzimático, denominado replissomo, que atua de maneira coordenada em cada réplicon, do qual participam proteínas específicas, como *helicases*, *proteínas de ligação ao DNA fita simples* e *topoisomerases*.

Inicialmente, as duas cadeias de nucleotídios que compõem a dupla hélice são desenroladas e separadas pela ação da enzima *DNA helicase*, que promove a quebra das pontes de hidrogênio entre as bases nitrogenadas, permitindo que cada fita seja replicada, isto é, sirva de molde para a síntese de um novo filamento complementar. Uma vez desenroladas, é necessário que as fitas de DNA permaneçam separadas na forma *unifilamentar*. A manutenção dos filamentos da dupla hélice afastados durante a replicação se dá pela imediata ligação de proteínas ao DNA unifilamentar – proteínas SSB (*single strand binding proteins*), que revestem cada uma das fitas individualmente, impedindo que elas retomem a forma de dupla hélice. As *SSB* são, portanto, proteínas desestabilizadoras da hélice, cuja função é manter as cadeias do DNA separadas. É importante observar que, em função da velocidade na qual a duplicação ocorre, tanto em procariotos quanto em eucariotos, a molécula de DNA deverá girar centenas de vezes por minuto para desenrolar a dupla hélice, sob ação das DNA helicases. Assim, à medida que o desenrolamento acelerado da hélice de DNA prossegue, ocorre um aumento de tensão na dupla fita, provocado pela super-helicoidização da molécula à frente da forquilha de replicação, que é solucionado pela ação das *DNA topoisomerases*. Essas enzimas catalisam cortes temporários em uma ou ambas as fitas da molécula juntamente com movimentos localizados que desfazem as torções da hélice, diminuindo a tensão e impedindo a superespiralização da molécula na região da forquilha de duplicação. As próprias DNA topoisomerases religam as quebras efetuadas.

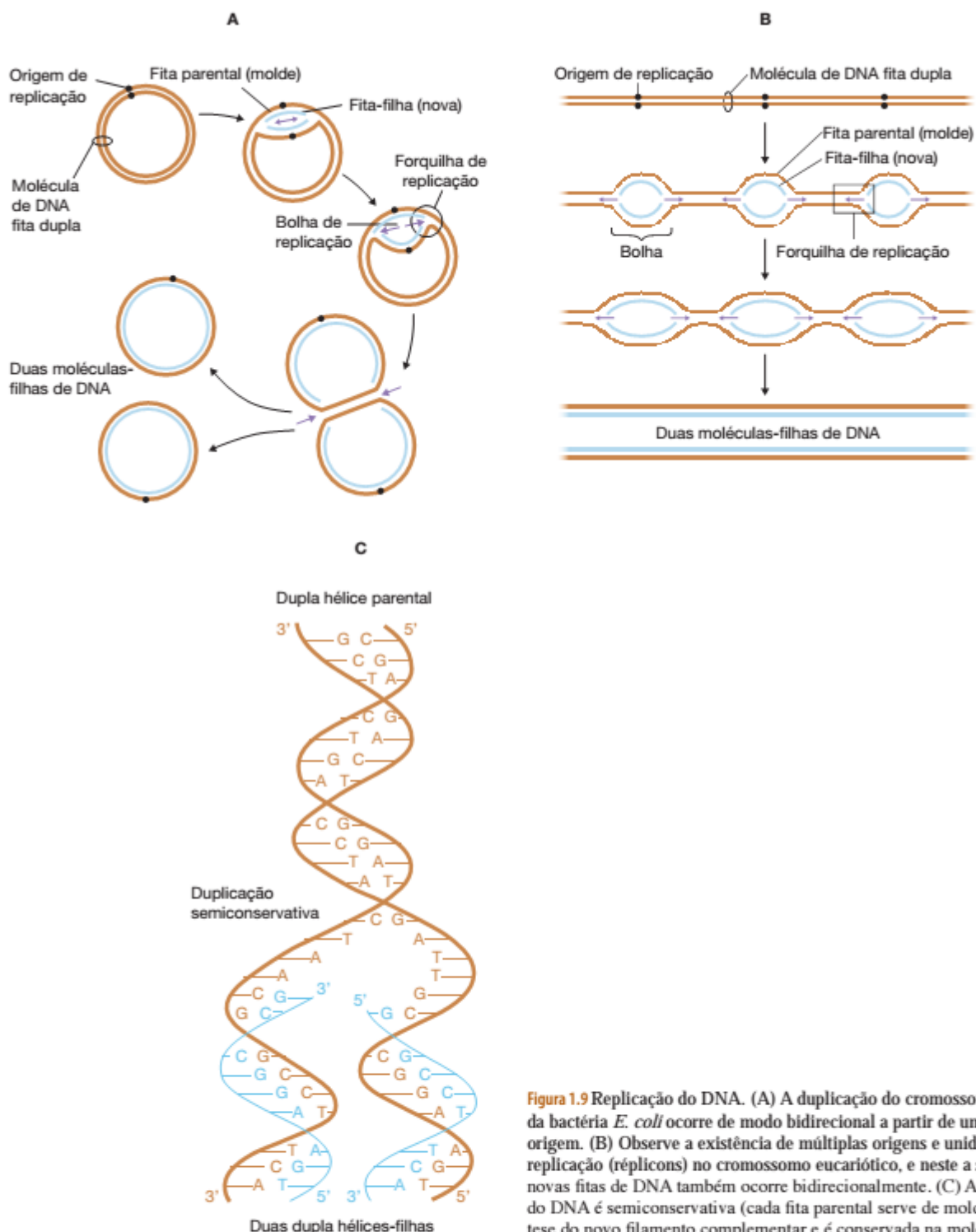



Figura 1.9 **Replicação do DNA. (A) A duplicação do cromossomo circular da bactéria *E. coli* ocorre de modo bidirecional a partir de uma única origem. (B) Observe a existência de múltiplas origens e unidades de replicação (réplicons) no cromossomo eucariótico, e neste a síntese das** novas fitas de DNA também ocorre bidirecionalmente. (C) A duplicação do DNA é semiconservativa (cada fita parental serve de molde para a síntese do novo filamento complementar e é conservada na molécula-filha).

DNA polimerase
Enzima responsável pela replicação do DNA. Todas as DNA polimerases replicativas adicionam desoxirribonucleotídios à fita em formação no sentido 5′ → 3′

Formação da cadeia polinucleotídica

Uma vez separados e estabilizados os filamentos da dupla hélice de DNA, inicia-se o processo, propriamente dito, de síntese das novas fitas complementares. Em todos os organismos vivos, **a replicação do DNA ocorre no sentido 5′ → 3′ pela ação da enzima DNA polimerase, que** catalisa a adição de desoxirribonucleotídios na cadeia em formação, tendo a fita parental como

molde, mantendo a complementaridade da sequência de bases da molécula original. A Figura 1.10 mostra um modelo geral da forquilha de replicação com os principais componentes que atuam na duplicação do DNA.

Existem múltiplas *DNA polimerases* em procariotos e eucariotos. Na bactéria *E. coli* são conhecidos cinco tipos (DNA polimerase I, II, III IV e V). A DNA polimerase III é uma enzima complexa formada por múltiplas subunidades, responsável pela replicação do DNA bacteriano e capaz de adicionar, aproximadamente, 500 nucleotídios por segundo, de modo a síntese de DNA nas bactérias ocorre em uma velocidade 10 vezes superior em relação à duplicação do DNA humano, em que, aproximadamente, 3.000 nucleotídios são adicionados por minuto. Em eucariotos, cerca de uma dezena de DNA polimerases são conhecidas, sendo sete delas bem caracterizadas em mamíferos (DNA polimerase α, β, γ, δ, ε, ζ e η). Nestes, três DNA polimerases (α, δ e ε) são essenciais à duplicação do DNA nuclear, e a DNA polimerase γ é responsável pela replicação do genoma mitocondrial. Todas as DNA polimerases replicativas existentes em procariotos e eucariotos são capazes de catalisar a síntese de novas cadeias de DNA somente na direção $5' \rightarrow 3'$, a partir da leitura de nucleotídios da fita parental de sentido $3' \rightarrow 5'$, e, embora sejam enzimas altamente precisas no que se refere à incorporação de nucleotídios na cadeia nascente, erros podem ocorrer, resultando em mutações, como veremos detalhadamente no Capítulo 3. Em média, a cada um bilhão de bases sintetizadas, um nucleotídio é incorretamente inserido. Considerando que a duplicação da molécula de DNA sem erros é de extrema importância para a sobrevivência do organismo, a célula tem mecanismos próprios de revisão e correção de falhas ocorridas durante a síntese. Durante a replicação, DNA polimerases com função de reparo, como a DNA polimerase III em bactérias, que além de atividade de polimerização também tem uma atividade de exonuclease, e as DNA polimerases β, ζ e η em mamíferos, identificam, removem e corrigem incorporações incorretas de nucleotídios nas cadeias em formação, à medida que elas são sintetizadas, garantindo a fidelidade do processo.

Uma característica comum a todas as DNA polimerases replicativas, tanto as de procariotos quanto as de eucariotos, é que nenhuma delas é capaz de iniciar a síntese de uma fita de DNA *de novo*, ou seja, tendo como molde apenas a fita parental. Todas elas, além de realizarem a

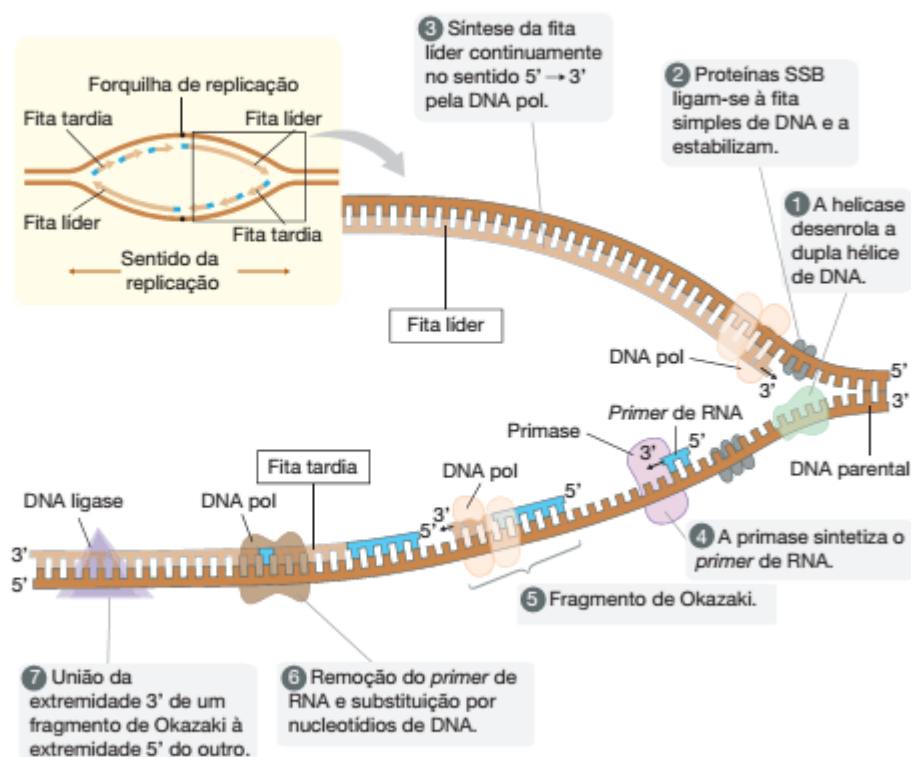



Figura 1.10 Representação de uma forquilha de replicação do DNA, mostrando as principais enzimas e proteínas que participam do processo de duplicação.

Primer (oligonucleotídio iniciador)

Pequeno segmento de RNA iniciador da replicação do DNA em procariotos e eucariotos que fornece a extremidade 3′-OH, à qual a DNA polimerase adiciona nucleotídios

Primase

Enzima responsável pela síntese do *primer* de RNA

polimerização de nucleotídios somente no sentido 5′→ 3′, necessitam de um *oligonucleotídio iniciador*, que forneça a extremidade 3′-OH livre, para que, a partir desta, possam adicionar os nucleotídios seguintes na fita em formação, de acordo com as bases presentes na fita parental. Essa extremidade 3′-hidroxila é fornecida pelo *primer de RNA*, um pequeno segmento iniciador, com apenas 10 ribonucleotídios de comprimento em eucariotos e 10 a 60 ribonucleotídios em procariotos. O *primer* de RNA é sintetizado por uma RNA polimerase específica denominada primase. As RNA polimerases, diferentemente das DNA polimerases, são enzimas capazes de sintetizar uma cadeia de RNA a partir de uma fita molde de DNA. Assim, a duplicação do DNA é iniciada pelo *primer* de RNA, que fornece a extremidade 3′-OH livre necessária para que a *DNA polimerase* possa proceder à polimerização de nucleotídios (Figura 1.11A). Antes do término da replicação, esses pequenos segmentos de RNA são removidos por enzimas específicas, que fazem a revisão dos nucleotídios incorporados durante a síntese, e são substituídos por DNA pela ação da DNA polimerase.

Devido à polaridade oposta dos filamentos que formam a dupla hélice e ao fato de as DNA polimerases replicativas atuarem apenas no sentido 5′→ 3′, a replicação da molécula de DNA ocorrerá distintamente nas duas fitas.

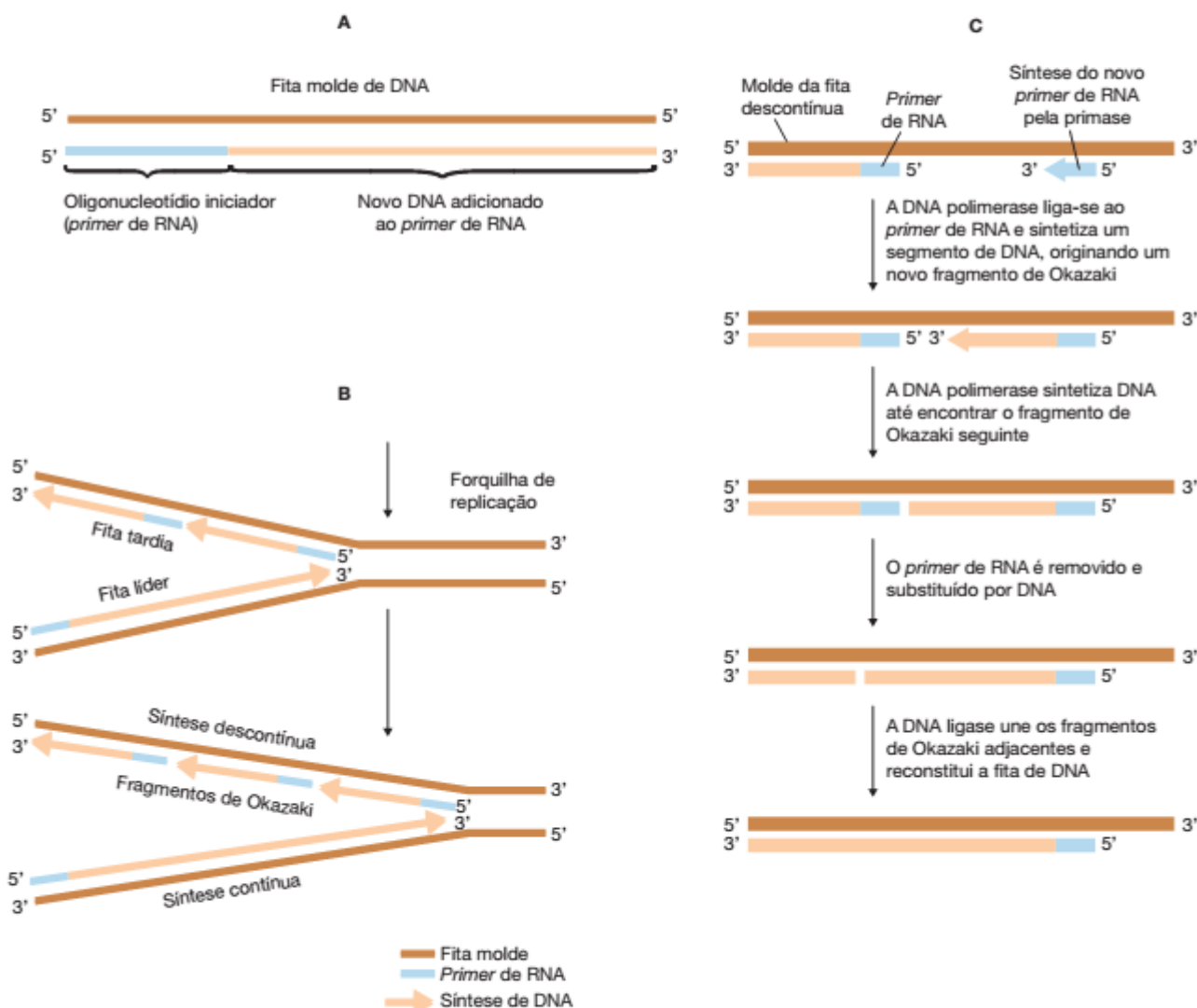



Figura 1.11 Início e progressão da síntese do DNA. (A) A enzima primase sintetiza, no sentido 5′→3′, um pequeno segmento iniciador de RNA (*primer*) complementar à fita molde, que fornece a extremidade 3′-OH para a DNA polimerase efetuar a polimerização dos nucleotídios. (B) Observe a progressão oposta da síntese de DNA na forquilha de replicação: na fita líder, a polimerização é contínua e na fita atrasada é descontínua, por meio da formação sequencial dos fragmentos de Okazaki. Note que os fragmentos mais próximos à forquilha são os que foram sintetizados mais recentemente. (C) Detalhe da fita descontínua, na qual os fragmentos de Okazaki formados são unidos, transformando a fita descontínua em um filamento contínuo.

Fita líder (*leading strand*)
Fita de DNA nascente polimerizada de modo contínuo, no mesmo sentido da abertura da forquilha de replicação

Fita atrasada ou tardia (*lagging strand*)
Fita de DNA sintetizada de modo descontínuo, na direção oposta à abertura da forquilha, por meio de curtos fragmentos denominados fragmentos de Okazaki

Fragmento de Okazaki
Curta sequência de nucleotídios sintetizada durante a replicação descontínua da fita tardia de DNA

Síntese semidescontínua
Característica da replicação do DNA. Por ser contínua em uma fita e descontínua na outra, a síntese é semidescontínua

DNA ligase
Enzima que catalisa a união de polinucleotídios. Ela estabelece a ligação covalente entre a extremidade 5′ de um fragmento de DNA e a extremidade 3′ de outro fragmento. Responsável pela união dos fragmentos de Okazaki

Na fita líder (*leading strand*), a síntese do novo filamento será *contínua*, ocorrendo na mesma direção da abertura da forquilha de replicação, e se dará a partir da formação de um *único primer de RNA*. Na outra fita, denominada fita atrasada ou tardia (*lagging strand*), a replicação deverá ocorrer no sentido oposto à abertura da forquilha e se dará por um processo diferente, por meio da síntese de *vários iniciadores de RNA (primers)* que serão formados à medida que a forquilha de replicação se abrir. Assim, diferentemente da fita líder, *na fita atrasada a duplicação do DNA ocorrerá de maneira descontínua*, pela síntese de curtos segmentos, com 100 a 200 nucleotídios de extensão em eucariotos e 1.000 a 2.000 nucleotídios em procariotos, constituídos por RNA e DNA, denominados fragmentos de Okazaki (Figura 1.11B). À medida que a polimerização da fita atrasada prosseguir, o *primer* de RNA de cada *fragmento de Okazaki* será removido e, em seguida, DNA polimerases sintetizarão um segmento de DNA a partir da extremidade 3′-OH livre do *fragmento de Okazaki* adjacente. Em seguida, a enzima DNA ligase efetuará a união dos *fragmentos de Okazaki* contíguos, reconstituindo a fita atrasada de DNA, que se tornará um filamento contínuo. Em função de a replicação do DNA ocorrer de modo ininterrupto na fita-líder e de maneira descontínua na fita atrasada, dizemos que a síntese de DNA é semidescontínua. (Figura 1.11C).

Embora a síntese da fita líder e da fita atrasada ocorra em sentidos físicos contrários, em função da estrutura da dupla fita de DNA, a polimerização de ambos os filamentos, contínuo e descontínuo, ocorre em uma mesma direção química (5′⟶ 3′).

Término da replicação nas extremidades dos cromossomos lineares de eucariotos: telômeros e telomerase

Entre as diferenças observadas na replicação do DNA de procariotos e eucariotos, talvez a mais marcante seja referente à finalização do processo, uma vez que a estrutura cromossômica interfere diretamente com o término da síntese. Ao contrário do cromossomo circular de bactérias, que, como mostrado na Figura 1.9A, não tem pontas e sua replicação progride normalmente até ser finalizada, a duplicação semiconservativa das extremidades dos cromossomos lineares eucarióticos representa um problema à parte, decorrente da impossibilidade de o aparato de síntese replicar a extremidade 5′ da fita atrasada. É importante observar que a dificuldade relativa à replicação das pontas dos cromossomos eucarióticos refere-se, especificamente, à extremidade da fita de DNA na qual a replicação ocorre de modo *descontínuo* (*fita atrasada ou tardia*), uma vez que, na fita líder, depois de iniciada a síntese de DNA, por esta ocorrer no mesmo sentido da abertura da forquilha, ela pode prosseguir continuamente até o final do filamento molde.

Como vimos, durante a replicação da fita descontínua, cada fragmento de Okazaki contém um *primer* de RNA, que é removido antes da união desses fragmentos, sendo o espaço que permanece preenchido por uma DNA polimerase, que adiciona nucleotídios ao grupo 3′-OH adjacente, recompondo a fita atrasada. Desta maneira, essa fita torna-se uma cadeia contínua. Entretanto, quando o último oligonucleotídio iniciador de RNA do fragmento de Okazaki da fita tardia é removido, não será possível reconstituir esse pequeno segmento linear terminal de DNA, em função de não existir uma extremidade 3′-hidroxila próxima necessária à ação enzimática de adição de desoxirribonucleotídios. Assim, a cada ciclo de replicação, um segmento de DNA terminal na fita atrasada, equivalente em extensão ao *primer* de RNA, não será duplicado, resultando no encurtamento da ponta de uma das fitas recém-sintetizadas. Ou seja, se essa dificuldade não for sanada, sempre que a dupla hélice linear de DNA duplicar, o resultado será a síntese de uma molécula com extremidades desiguais em função de uma das fitas sintetizadas (fita descontínua) ser mais curta. Observando a Figura 1.12A pode-se compreender melhor o que ocorre.

Na verdade, a síntese de DNA na região terminal da fita tardia só é possível em função da composição diferenciada das *regiões teloméricas* e da existência de uma enzima particular chamada *telomerase*. Relembrando, as extremidades dos cromossomos lineares de eucariotos são protegidas por estruturas específicas, os *telômeros*, formados por inúmeras sequências curtas de DNA não codificante, repetidas em tandem. As unidades de repetições que formam o DNA telomérico têm pequenas variações de sequência entre eucariotos, e esta é idêntica

em todos os vertebrados, incluindo os seres humanos (5′ TTAGGG 3′). Em função de sua composição de bases, a fita com a terminação 3′ denomina-se fita rica em guaninas (G), e a fita complementar, com terminação 5′, é denominada fita rica em C. Desta forma, a impossibilidade de síntese das pontas da fita tardia, a cada ciclo de replicação, acarretaria a perda de algumas unidades de repetições que formam os telômeros. Após inúmeras divisões celulares, a perda progressiva de segmentos terminais na fita de DNA, em potencial, poderia se estender além das regiões teloméricas e atingir as sequências codificadoras, ocasionando a perda de informações genéticas essenciais para a célula.

De fato, o segmento de DNA terminal na fita atrasada, correspondente ao último *primer* de RNA, pode ser reconstituído pela ação da enzima *telomerase*, específica de eucariotos. A *telomerase* é uma enzima diferenciada que contém uma *pequena molécula de RNA* (cuja sequência é complementar à região terminal da fita rica em G), que serve de molde para a extensão dos telômeros. Ou seja, essa ribonucleoproteína é capaz de adicionar pontas aos cromossomos lineares de eucariotos por meio de um processo enzimático, conforme ilustrado na Figura 1.12B.

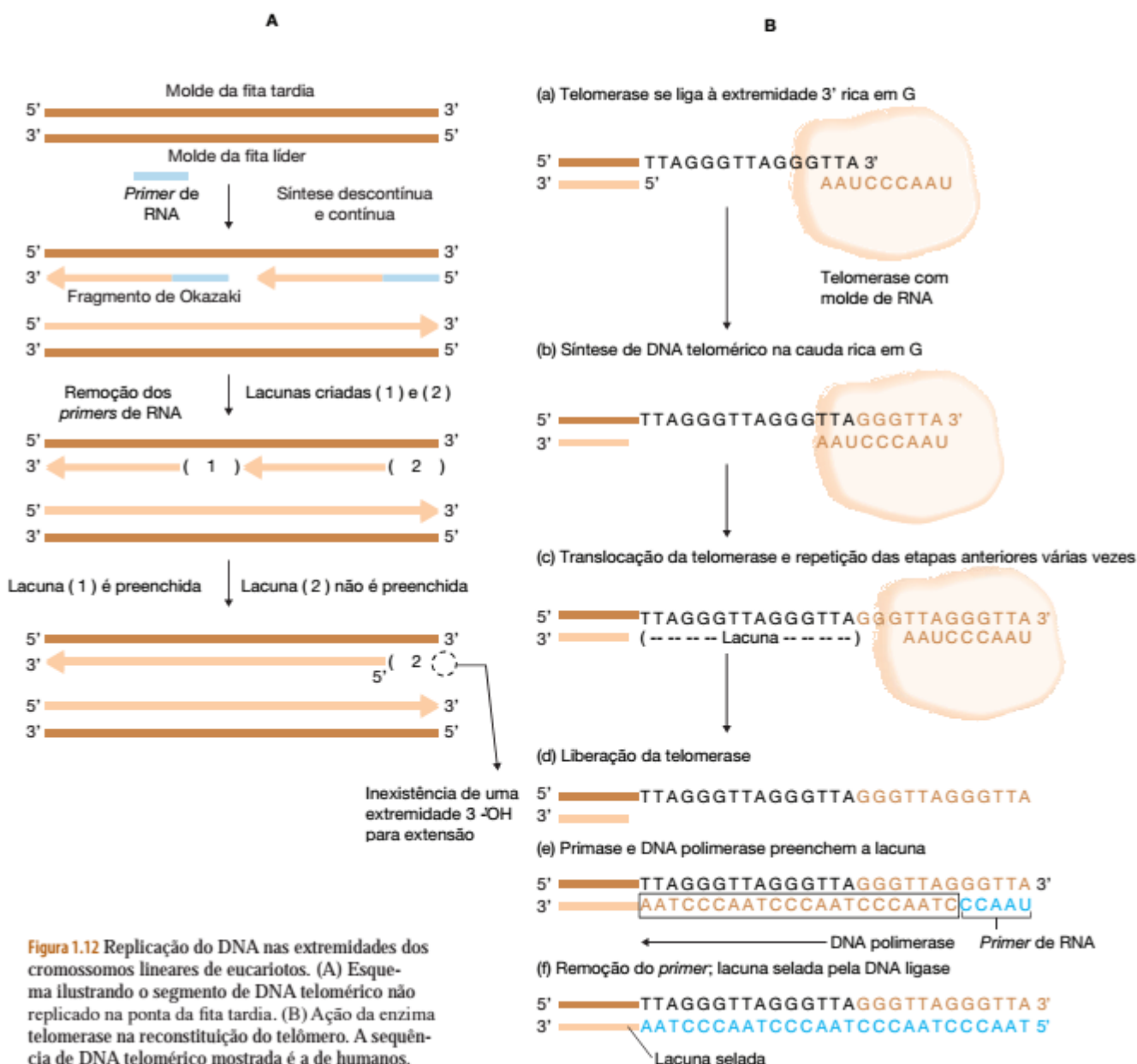



Figura 1.12 Replicação do DNA nas extremidades dos cromossomos lineares de eucariotos. (A) Esquema ilustrando o segmento de DNA telomérico não replicado na ponta da fita tardia. (B) Ação da enzima telomerase na reconstituição do telômero. A sequência de DNA telomérico mostrada é a de humanos.

Transcrição reversa
Síntese de DNA a partir de um molde de RNA

A molécula de RNA presente na telomerase torna possível que ela reconheça e se ligue à extremidade 3′ do DNA telomérico, rica em G, que se projeta. Por meio de transcrição reversa, o segmento de RNA da telomerase servirá de molde para a síntese de seu DNA complementar, e ela adicionará uma unidade de repetição a essa extremidade, deslocando-se em seguida, de modo a sintetizar novas unidades de DNA telomérico, uma por vez, estendendo a fita rica em G na direção 5′→ 3′. Observe que a telomerase não efetua a síntese do segmento que falta na fita rica em C oposta (com a terminação 5′), atuando apenas como um molde para a extensão da ponta 3′ do filamento de DNA rico em G. Após alguns ciclos de extensão dessa fita, a telomerase será liberada. Segue-se, então, a síntese do *primer* de RNA, pela *DNA primase*, na extremidade 5′ da fita rica em C, e a síntese de DNA no espaço existente pela *DNA polimerase*.

A adição de segmentos de DNA telomérico ao prolongamento 3′ pela telomerase possibilita que as enzimas DNA primase e DNA polimerase possam usar essa extensão como molde para sintetizar o segmento de DNA complementar, recompondo a extremidade da fita tardia.

Nas células humanas com alta capacidade proliferativa, como as embrionárias e as da linhagem germinativa, o comprimento cromossômico é mantido pela telomerase, que se encontra ativa nessas células. De maneira semelhante, inúmeras células cancerosas também mantêm a atividade dessa enzima, o que lhes confere capacidade de proliferação infinita.

Por sua vez, a maior parte das células somáticas de mamíferos (incluindo humanos) sintetiza muito pouca ou nenhuma telomerase, de maneira que, a cada ciclo de replicação do DNA, os telômeros se tornam menores. A ausência de atividade dessa enzima nas células somáticas ocasiona limitação em sua capacidade de proliferação (*senescência replicativa*). O cultivo *in vitro* de células humanas mostra que estas se dividem por um número específico de gerações (geralmente 20 a 70) antes que ocorra a senescência e a morte. É possível observar em culturas de células somáticas uma correlação direta entre o tamanho dos telômeros e o tempo de sobrevivência da célula (células com cromossomos que apresentam telômeros mais curtos sobrevivem menos). O progressivo encurtamento das extremidades cromossômicas ocorrido durante a duplicação do DNA nas células que não sofrem ação da telomerase, na verdade, funciona como um relógio biológico e é parte do processo natural de envelhecimento celular. Outra forte evidência da relação entre o tamanho dos telômeros e o envelhecimento advém de algumas doenças humanas caracterizadas por envelhecimento prematuro. As *progerias*, como são conhecidas, constituem um grupo de doenças genéticas raras (síndrome de Werner, disqueratose congênita, síndrome de Hutchinson-Gilford) que cursam com um processo acelerado de envelhecimento, o qual, nas formas mais graves, inicia-se logo em seguida ao nascimento, levando à morte na adolescência. O envelhecimento precoce observado nos pacientes com progeria correlaciona-se com a baixa capacidade de proliferação das células somáticas desses indivíduos, em decorrência do encurtamento de seus telômeros.

A telomerase compensa a perda de segmentos de DNA telomérico, que ocorre durante a replicação do DNA, e sua atividade é relacionada com a reconstituição dos telômeros de eucariotos. Na ausência desta enzima, as extremidades cromossômicas tornam-se progressivamente menores.

Na Tabela 1.1, encontram-se resumidas as principais características da replicação do DNA em procariotos e eucariotos.

Tabela 1.1 Características gerais da duplicação do DNA.

Eucariotos	Procariotos
Ocorre na fase S do ciclo celular	Ocorre continuamente
Semiconservativa	Semiconservativa
Semidescontínua	Semidescontínua
Cromossomos com múltiplas origens de replicação (réplicons)	Cromossomo com uma única origem de replicação (um réplicon)
Bidirecional	Bidirecional
Síntese no sentido 5′→ 3′	Síntese no sentido 5′→ 3′
Necessidade de oligonucleotídios iniciadores (*primer* de RNA)	Necessidade de oligonucleotídios iniciadores (*primer* de RNA)
Formação de fragmentos de Okazaki	Formação de fragmentos de Okazaki
Replicação das moléculas lineares de DNA representa um problema na extremidade da fita descontínua	Cromossomo circular replica sem problemas

RESUMO

- Na maioria dos organismos, a informação genética está armazenada no ácido desoxirribonucleico (DNA). Em alguns vírus, ela está contida no ácido ribonucleico (RNA)
- Nas células eucarióticas, a maior parte do DNA está no núcleo, embora também exista DNA em algumas organelas, como as mitocôndrias e os cloroplastos
- Os ácidos nucleicos (DNA e RNA) são longos polímeros formados por nucleotídios, e estes são constituídos por: açúcar (desoxirribose ou ribose), fosfato e bases nitrogenadas (pirimidinas e purinas)
- As moléculas de RNA são curtas e unifilamentares. O DNA é uma dupla hélice helicoidal, formada por cadeias de polinucleotídios antiparalelas e complementares
- A maneira como os nucleotídios ligam-se uns aos outros confere uma polaridade às fitas do DNA e do RNA (extremidade 5′ – presença de um grupo fosfato livre, e extremidade 3′ – presença de um grupo hidroxila livre)
- A dupla hélice de DNA é uma estrutura dinâmica, que pode assumir diferentes conformações (formas A, B e Z). A forma normal da molécula é conhecida como DNA B
- O termo *genoma* refere-se ao conteúdo total de DNA presente na célula de um organismo
- O genoma eucariótico é formado por DNA de cópia única e DNA repetitivo (sequências moderadamente ou altamente repetitivas)
- O DNA altamente repetitivo constitui cerca de 5% do genoma humano e é representado pelo DNA satélite ou alfoide, presente nas regiões centroméricas. O DNA moderadamente repetitivo representa 45% do genoma e inclui os minissatélites, os microssatélites e os elementos de transposição (SINE e LINE)
- O genoma nuclear humano representa mais de 99,95% do DNA celular e encontra-se distribuído em 46 cromossomos lineares (22 pares de autossomos e 2 cromossomos sexuais X e Y), cada um deles constituído por uma única molécula de DNA associada a proteínas específicas formando a cromatina
- O genoma mitocondrial humano é extremamente compacto, formado por uma pequena molécula de DNA circular dupla fita, com 16.569 pb, que contém 37 genes
- Menos de 2% do DNA nuclear humano são sequências gênicas codificadoras de proteínas
- Os genes nucleares humanos, como os da maioria dos eucariotos, são formados por regiões codificadoras (éxons), intercaladas por regiões não codificadoras (íntrons), e extremamente variáveis em tamanho
- A associação do DNA a proteínas histonas possibilita que a cromatina eucariótica apresente diferentes níveis de compactação, desde a sua organização nos nucleossomos até a formação do solenoide e do cromossomo mitótico
- A heterocromatina refere-se à cromatina altamente condensada e transcricionalmente inativa. A eucromatina representa a forma menos compactada da cromatina
- O ciclo celular engloba dois estágios: o período de interfase ou não divisão, constituído pelas fases G_1, S (síntese do DNA) e G_2, e a divisão celular propriamente dita (fase M), que engloba a mitose e a citocinese
- A duplicação do DNA tem início em uma região específica da molécula denominada origem de replicação. Uma única origem de replicação (réplicon) é observada no cromossomo bacteriano, enquanto os cromossomos eucarióticos têm inúmeros réplicons
- A replicação do DNA em procariotos e eucariotos é bidirecional, semiconservativa e semidescontínua
- Um complexo sistema proteico-enzimático, do qual participam helicases, proteínas de ligação ao DNA fita simples e topoisomerases, está envolvido no processo de síntese do DNA. As helicases desenrolam a dupla hélice; as proteínas SSB mantêm as fitas separadas; e as topisomerases controlam a tensão provocada pelo desenrolamento da molécula
- Em todos os organismos, a replicação do DNA ocorre no sentido 5′→ 3′ pela ação da enzima DNA polimerase, que adiciona desoxirribonucleotídios na cadeia em formação, complementares à molécula parental
- As DNA polimerases replicativas necessitam de um *primer* de RNA, que forneça a extremidade 3′-OH livre, para que possam adicionar nucleotídios à fita em formação
- A polaridade oposta dos filamentos que formam a dupla hélice e o fato de as DNA polimerases efetuarem a síntese no sentido 5′→ 3′ implica que a replicação da molécula ocorrerá de modo contínuo em uma das duas fitas (líder) e de modo descontínuo na outra (atrasada ou tardia), pela formação de curtos segmentos denominados fragmentos de Okazaki
- A duplicação do segmento de DNA terminal na fita atrasada dos cromossomos lineares eucarióticos representa uma dificuldade. Este segmento pode ser sintetizado por transcrição reversa pela enzima telomerase, que promove a extensão dos telômeros
- A existência de pouca ou nenhuma atividade da telomerase na maioria das células somáticas de mamíferos (incluindo humanos) implica que após um determinado número de divisões elas entrarão em senescência replicativa.

AUTOAVALIAÇÃO

1.1 Que são ácidos nucleicos e qual a sua composição química?
1.2 Quais as principais diferenças entre as moléculas de DNA e de RNA?
1.3 Diferencie nucleotídio de nucleosídio.
1.4 Sabendo-se que o conteúdo de G em uma molécula de DNA é 30%, qual o percentual das demais bases nitrogenadas (C, A e T)?
1.5 Em relação às conformações assumidas pela dupla hélice de DNA, diferencie a forma B das formas A e Z.
1.6 Comente: "A forma B do DNA é uma dupla hélice dextrógira".
1.7 Quais as principais diferenças entre o genoma nuclear e o genoma mitocondrial humano?
1.8 Defina cromatina.
1.9 Por que se diz que os cromossomos lineares de eucariotos são uninêmicos?

1.10 Em que consiste o DNA satélite?
1.11 Defina os tipos de sequências, em relação ao número de cópias, encontradas nos genomas de eucariotos.
1.12 Que são elementos de transposição? Conceitue LINE e SINE.
1.13 Diferencie microssatélite de minissatélite.
1.14 Por que os VNTR e os STR são amplamente usados como marcadores moleculares?
1.15 Quantos genes humanos codificadores de proteínas existem a partir das estimativas do projeto genoma humano?
1.16 Como explicar que menos de 2% do DNA nuclear humano codifique proteínas, considerando a grande diversidade das proteínas humanas conhecidas?
1.17 Diferencie éxon de íntron.
1.18 Qual a função das proteínas histonas?
1.19 Conceitue nucleossomo e descreva sua estrutura.
1.20 O que é solenoide?
1.21 Diferencie heterocromatina de eucromatina.
1.22 Conceitue as diferentes fases do ciclo celular de uma célula eucariótica.
1.23 Em que fase de seu ciclo de vida a célula passa a maior parte?
1.24 Explique como se inicia a síntese do DNA em procariotos e eucariotos.
1.25 Defina réplicon.
1.26 Qual a função das enzimas helicases, topoisomerases e polimerases?
1.27 Como é possível, após a abertura da dupla hélice de DNA, manter as fitas separadas durante o processo de replicação?
1.28 Comente a afirmativa: "A duplicação do DNA é bidirecional."
1.29 Por que a duplicação do DNA é semiconservativa?
1.30 Por que a síntese de DNA é contínua em uma fita (*leading strand*) e descontínua na outra (*lagging strand*)?
1.31 O que é *primer* e por que ele é formado?
1.32 Que são fragmentos de Okazaki?
1.33 O que é telômero e qual a sua função?
1.34 Explique a principal característica e função da enzima telomerase?
1.35 Qual a dificuldade para replicar o DNA nas extremidades dos cromossomos lineares de eucariotos?
1.36 O que você entende por senescência replicativa?
1.37 Defina progeria.
1.38 A partir do que foi apresentado, você seria capaz de elaborar um resumo explicando todo o processo de replicação do DNA, salientando as principais diferenças observadas entre procariotos e eucariotos?

BIBLIOGRAFIA

ALBERTS, B.; JOHNSON, A.; LEWIS, J. *et al. Biologia molecular da célula.* 5ª ed., Porto Alegre: Artmed, 2010.

BANG, M-L.; CENTNER, T.; FORNOFF, F. *et al.* The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. 2001; *Circ Res* 89: 1065-1072.

BENNET, R.L.; STEINHAUS, K.H.; UHRICH, S.B. *et al.* 1995. Recommendation for standardized human pedigree nomenclature. *Am J Hum Genet* 56: 745-752.

CAMPBELL, N.A.; REECE, J.B. *Biologia.* 8ª ed., Porto Alegre: Artmed, 2010.

CLAMP, M.; FRY, B.; KAMAL, M. *et al.* Distinguishing protein-coding and noncoding genes in the human genome. *PNAS* 2007; 104: 19428-19433.

DEN DUNNEN, J.T.; GROOTSCHOLTEN, R.M.; BAKKER, E. *et al.* Topography of the Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Hum Genet* 1989; 45: 835-847.

ENSEMBL v. 65 (Dezembro de 2011). Disponível em: http://www.ensembl.org/Homo_sapiens/Info/Index

GRIFFITHS, A.J.F.; WESSLER, S.R.; LEWONTIN, R.C. *et al. Introdução à genética.* 9ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860-921.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. Finishing the euchromatic sequence of the human genome. *Nature* 2004; 431: 931-945.

KELLER, E.F. *O século do gene.* 1ª ed., Belo Horizonte: Crisálida, 2002.

KLUG, W.S.; CUMMINGS, M.R.; SPENCER, C.A. *et al. Conceitos de genética.* 9ª ed., Porto Alegre: Artmed, 2010.

LEWIN, B. *Genes IX.* 9ª ed., Porto Alegre: Artmed, 2009.

NASSEH, I.E.; TENGAN, C.H.; KIYOMOTO, B.H. *et al.* Doenças mitocondriais. *Rev Neurociências* 2001; 9: 60-69.

ONLINE MENDELIAN IN MAN (OMIM). Disponível em: http://www.ncbi.nlm.nih.gov/omim/. Acesso em 01/02/2012.

ROSS, M.T.; GRAFHAM, D.V.; COFFEY, A.J. *et al.* The DNA sequence of the human X chromosome. *Nature* 2005; 434: 325-337.

SKALETSKY, H.; KURODA-KAWAGUCHI, T.; MINX, P.J. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 2003; 423: 825-837.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

STRACHAN, T.; READ, A.P. *Genética molecular humana.* 2ª ed., Porto Alegre: Artmed, 2002.

VENTER, J.C.; ADAMS, M.D.; MYERS, E.W. *et al.* The sequence of the human genome. *Science* 2001; 291: 1304-1351.

WALTER, J.; PAULSEN, M. Imprinting and disease. *Semin Cell Dev Biol* 2003; 14: 101-110.

WILLARD, H.; GINSBURG, G. *Genomic and personalized medicine*, 1ª ed., EUA: Academic Press, 2008.

2

Genes e Controle da Expressão Gênica

Objetivos de estudo

- Saber caracterizar os principais elementos da estrutura de um gene funcional
- Compreender as diferentes etapas que compõem o processo de transcrição gênica em procariotos e eucariotos
- Reconhecer os diferentes tipos de RNA produzidos na transcrição, suas funções celulares e seus processamentos
- Entender como o processo de tradução acontece em procariotos e eucariotos, identificando suas peculiaridades
- Compreender como os genes procarióticos e eucarióticos podem ser regulados nos diferentes níveis
- Ser capaz de redefinir o conceito clássico de "gene"

Conceitos-chave do capítulo

Aminoácidos
Aminoacilação
Aminoacil-tRNA sintetase
Amplificação gênica
Anticódon
CAAT boxe
Cadeia polipeptídica
Capeamento 5′
Cauda poli(A)
Chaperonas
Código genético
Códon
Controle negativo
Controle positivo
Dicer
Dissomia uniparental
Dogma Central
Domínios de cromatina
Edição do RNA
Efeito de longo alcance
Elementos cisatuantes
Elementos de resposta
Epigenética
Estrutura primária
Estrutura quaternária
Estrutura secundária
Estrutura terciária
Exclusão de éxons
Éxon
Fase aberta de leitura
Fatores de transcrição transatuantes
Genes de choque térmico
Genes de manutenção
Heterodissomia
Holoenzima
Imprinting genômico
Imprinting parental
Inativação do X
Insuladores
Interferência por RNA
Íntron
Isodissomia
MBP
Metilação do DNA
miRNA
Modificação de histonas
Modificações pós-traducionais
mRNA
Mundo de RNA
Óperon
Óperon indutível
Óperon repressível
piRNA
Poliadenilação 3′
Polipeptídio
Poliribossomo
Posição de oscilação
Pré-mRNA
Pribnow boxe
Promotor
Pseudogene
Recomposição alternativa
Recomposição dos éxons
Reforçadores
Regulação espacial
Regulação temporal
Remodelagem de cromatina
Repressão catabólica
Ribossomo
Ribozimas
RISC
RITS
RNA antissenso
RNA policistrônico
rRNA
scRNA
Sequência consenso
Sequência de Shine-Dlagarno
Silenciadores
Silenciamento gênico
siRNA
Slicer activity
snoRNA
snRNA
snRNP
Spliceossomo
tmRNA
Tradução
Transcrição
Transcrição reversa
Transcrito primário
tRNA
Unidade de transcrição
Vigilância do mRNA

■ Introdução

Neste capítulo, abordaremos como a informação biológica de um gene se torna disponível para ser utilizada pela célula, processo conhecido como *expressão gênica*, e como esse processo é regulado em diferentes níveis. O genoma humano tem todas as instruções biológicas necessárias para o funcionamento de nosso organismo. Conforme ilustrado na Figura 2.1, os genes perpetuam a informação genética por meio da replicação do DNA, e a transcrição e a tradução são os dois mecanismos que a célula utiliza para expressar a informação biológica contida nos genes.

Transcrição
Síntese de uma molécula de RNA com base em um molde de DNA pela RNA polimerase

Tradução
Síntese de um polipeptídio com base em uma molécula de RNA mensageiro

Código genético
Conjunto de regras que governa a tradução de uma molécula de mRNA em uma proteína

Esses três processos baseiam-se na complementaridade de nucleotídios.

O Dogma Central da Biologia Molecular, elaborado por Francis Crick em 1958, durante muito tempo adotou o conceito de que todo gene era transcrito *unidirecionalmente* em uma molécula unifilamentar de RNA (*transcrição*), que, por sua vez, era obrigatoriamente traduzida em uma única proteína (*tradução*). O termo transcrição é utilizado porque a passagem da informação genética do DNA para o RNA envolve um fluxo de informações na linguagem de ácidos nucleicos, ao passo que, na tradução, a informação genética deve ser decodificada de uma linguagem de nucleotídios para uma linguagem de aminoácidos, de acordo com os preceitos do código genético.

Ao considerar que o fluxo da informação genética segue do DNA para o RNA e, posteriormente, para as proteínas, e que, assim como a replicação, todo o processo de expressão gênica precisa de proteínas, você pode pensar: O que surgiu primeiro: as proteínas ou os ácidos nucleicos? Embora haja controvérsias, indícios demonstram que o RNA surgiu antes do DNA, em um momento longínquo da evolução chamado de *Mundo de RNA*, como o veículo predominante de propagação da informação gênica. Essa hipótese foi traçada após a descoberta das moléculas de RNA, as ribozimas (en*zima* de ácido *ribo*nucleico), as quais apresentam capacidade autocatalítica de realizar suas próprias reações químicas e, por isso, representariam fósseis moleculares da origem da vida. Além da função de estocagem de informação genética, a otimização da habilidade catalítica desses RNA teria possibilitado a capacidade de sintetizar proteínas, as quais passaram a atuar como os grandes componentes estruturais e catalisadores das células. Assim, em virtude da maior estabilidade química, o papel de estocagem e replicação da informação genética foi atribuído ao DNA. Contudo, o RNA continuou a exercer funções essenciais para a célula, atuando inclusive na replicação do DNA e como mensageiro da informação genética do DNA até as proteínas.

Com o avanço do conhecimento molecular, ficou claro que existem exceções ao Dogma Central proposto por Crick.

Em primeiro lugar, em determinadas situações, o processo de transcrição do DNA para uma molécula de RNA pode não ocorrer em uma via de mão única. Retrovírus, como o HIV, e alguns

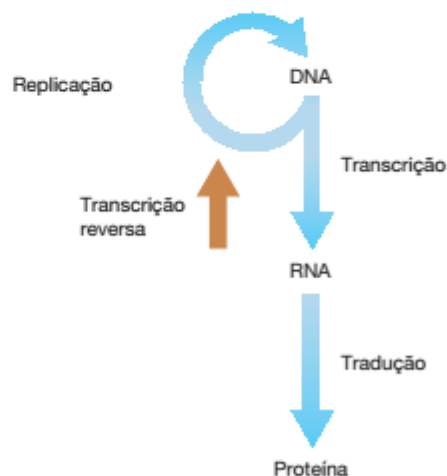



Figura 2.1 Caminhos que levam à expressão de um gene. A replicação do DNA é utilizada para perpetuar a informação genética, enquanto que a transcrição e a tradução permitem a expressão da informação biológica contida nos genes.

Transcrição reversa
Síntese de uma molécula de DNA com base em um molde de RNA, comum em retrovírus

cDNA
DNA complementar formado a partir de uma sequência de RNA

elementos de transposição têm a habilidade de transferir a informação genética contida em uma molécula de RNA para uma molécula de DNA apta a se integrar no genoma hospedeiro. A conversão do RNA para o DNA, chamada de transcrição reversa, acontece por causa da atuação de uma enzima *transcriptase reversa* que tem grandes aplicações em ensaios de Biologia Molecular. Em virtude da maior suscetibilidade química do RNA à degradação, causada pela presença de uma hidroxila livre no carbono 2′ da ribose, a transcriptase reversa possibilita que essas moléculas sejam convertidas *in vitro* em moléculas estáveis de DNA complementar (cDNA), mais fáceis de serem manipuladas em procedimentos laboratoriais.

Além do fluxo da informação genética poder voltar do RNA para o DNA, há outras exceções ao Dogma Central. Por exemplo, em alguns genes, a molécula de RNA é o produto gênico final, não sendo posteriormente traduzida em uma proteína; para outros genes, mais de uma proteína correlata pode ser traduzida com base em um único transcrito de RNA.

Estrutura gênica

Antes da apresentação dos conceitos relacionados com a expressão gênica propriamente dita e à sua regulação, mostraremos a organização interna dos genes humanos, a qual viabiliza sua expressão.

Embora o termo "gene" tenha sido citado pela primeira vez em 1909 por Wilhelm Johansen, Mendel, em sua publicação de 1866, já havia feito referência a este conceito usando a expressão "fator unitário" ou "caracter" como uma unidade herdada capaz de controlar um aspecto específico do fenótipo, exemplificado pela cor das ervilhas. Desde então, o conceito de gene tem mudado consideravelmente e, na Genética Moderna, passou a representar uma unidade funcional que controla a síntese de um produto final, o qual pode ser uma molécula de RNA ou um polipeptídio (proteína). Desse modo, podemos diferenciar dois tipos de genes funcionais no genoma humano: os que codificam moléculas de RNA e os que codificam proteínas.

Os genes que expressam RNA, e que estão envolvidos no controle da expressão gênica, inclusive a síntese de proteínas, representam de 2% a 5% do total.

Como citado anteriormente, esses genes passam apenas por um estágio da expressão gênica, a transcrição. O número de tipos de RNA produzidos nessa etapa é crescente e suas funções serão discutidas mais adiante. São eles: o RNA ribossômico (*rRNA*), o RNA transportador (*tRNA*), pequenos RNA nucleares (*snRNA*), pequenos RNA nucleolares (*snoRNA*), pequenos RNA citoplasmáticos (*scRNA*), microRNAs (*miRNA*) e pequenos RNA de interferência (*siRNA*). Por outro lado, os genes que codificam proteínas representam a maioria e exibem uma vasta diversidade em organização e tamanho. Sua expressão envolve dois passos. Primeiramente, a sequência de DNA de fita dupla do gene é transcrita em uma molécula unifilamentar de RNA denominada *RNA mensageiro* (mRNA). Em seguida, essa molécula intermediária de RNA é decodificada em uma sequência de aminoácidos que será moldada para formar o polipeptídio final. Desse modo, todos os genes eucariotos sofrem transcrição, mas nem todos são traduzidos em proteínas. Outro detalhe importante é que o fluxo de informações na tradução não é reversível como pode ocorrer na transcrição reversa.

Unidade de transcrição
Segmento de DNA que é transcrito em uma molécula primária de RNA e suas regiões reguladoras

As moléculas de DNA de nossas células abrangem informações genéticas que precisam estar dispostas de maneira organizada para que a expressão de um gene ocorra corretamente. Um segmento de DNA transcrito em uma molécula unifilamentar de RNA e as sequências necessárias para sua transcrição são chamados de unidade de transcrição.

A unidade de transcrição que codifica a proteína contém não só a sequência que será eventualmente traduzida em uma proteína (sequência codificante), mas também sequências reguladoras que direcionam e regulam a síntese da proteína. Uma unidade de transcrição pode conter um gene individual ou, mais raramente, alguns genes contínuos.

Sequências de DNA consenso
Sequências de nucleotídios comumente presentes em elementos genéticos do DNA ou RNA e que têm função específica

Cada unidade de transcrição é resultado de três elementos principais: um promotor, a sequência a ser efetivamente transcrita e um finalizador (Figura 2.2).

É essencial que a transcrição tanto de procariotos quanto de eucariotos seja iniciada em posições exatas determinadas por sequências de DNA consenso evolutivamente conservadas, as quais

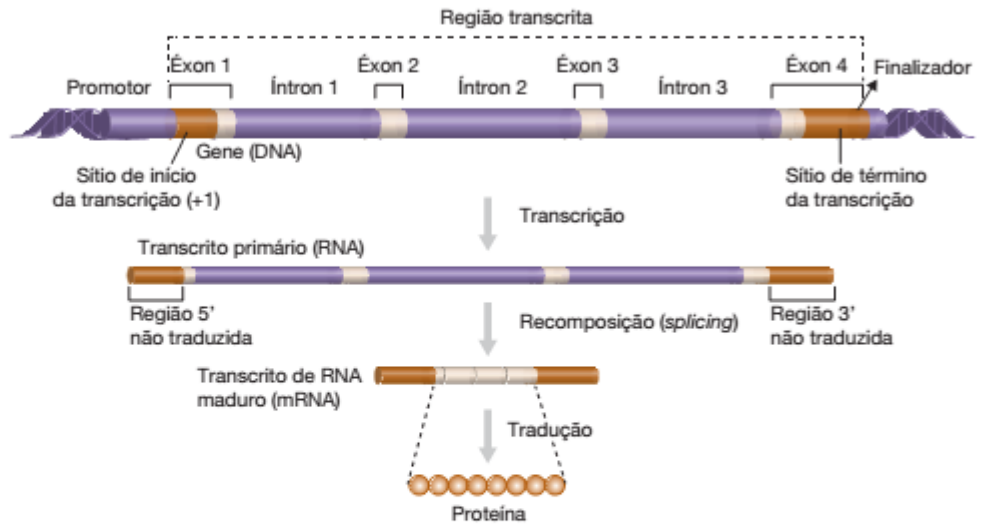



Figura 2.2 Estrutura clássica de um gene, destacando o promotor, a sequência efetivamente transcrita e o finalizador.

Promotor
Sequências conservadas, às quais a RNA polimerase se liga para iniciar a transcrição de um gene

são conhecidas como promotor. Em geral, o promotor situa-se antes do primeiro nucleotídio a ser transcrito, mas ele mesmo não é transcrito. Sua função é ser reconhecido pelo aparato proteico envolvido na transcrição, determinando o ponto exato de início do processo. O primeiro nucleotídio a ser transcrito recebe a numeração +1. Nucleotídios antecedentes a esse ponto recebem numeração negativa, ao passo que os nucleotídios posteriores a +1 recebem numeração positiva.

Além do promotor, existem sequências de DNA conservadas que delimitam o término da transcrição. Esses elementos finalizadores costumam fazer parte da sequência a ser transcrita, diferentemente do promotor, mas a transcrição só termina após o finalizador ter sido copiado. Entre o promotor e o finalizador está a sequência a ser efetivamente transcrita em uma molécula de RNA e, eventualmente, traduzida em proteína. Nesse segmento transcrito, para a maioria dos genes de eucariotos, existem sequências adicionais de DNA que interrompem o segmento que codificará uma proteína. Durante a expressão gênica, essas sequências de nucleotídios são removidas para gerar um transcrito de RNA maduro e que conterá apenas a sequência correspondente à proteína a ser sintetizada.

Éxon
Segmento de DNA que está representado no mRNA maduro e que participará como molde na síntese proteica

Íntron
Segmento de DNA que é transcrito, mas posteriormente removido do mRNA gerado, e que não servirá de molde na síntese proteica

Desse modo, vários genes em eucariotos são descontínuos e formados por regiões codificadoras de proteínas (éxons, de sequências *ex*pressas) e regiões intervenientes não codificadoras (íntrons, de sequências *int*ervenientes). Isso significa que todos os éxons e íntrons são inicialmente transcritos em um RNA. Após a transcrição, os íntrons são removidos, ao passo que os éxons são unidos na mesma ordem presente na molécula de DNA molde por recomposição, para gerar o mRNA maduro.

Pelo fato de os éxons participarem diretamente da formação das proteínas e estarem, portanto, sob maior pressão seletiva, regiões exônicas ao longo do genoma costumam ser muito mais conservadas intra e inter espécies do que os íntrons. Dessa forma, podemos dizer que os íntrons evoluem muito mais rapidamente do que os éxons. Os íntrons são raros em genes de bactérias.

O tamanho dos genes humanos costuma ser ditado pelo tamanho dos íntrons, que podem chegar a dezenas de milhares de bases; já os éxons costumam ser bem menores, com um tamanho médio de 200 pb. Assim, genes grandes costumam ter íntrons muito extensos. Os menores genes humanos conhecidos são os que codificam as proteínas histonas, e apresentam cerca de 500 pb ininterruptos (sem íntrons). Por outro lado, o gene que codifica a proteína distrofina, ausente ou não funcional na distrofia muscular, é considerado o maior gene humano, com cerca de 2,4 milhões de pares de bases. Cerca de 99% desse gene corresponde a seus 79 íntrons, o que faz com que ele demore cerca de 16 h para ser transcrito. O tamanho e o número de íntrons também parecem ter relação com a complexidade do organismo.

Transcrição

A transcrição é o primeiro estágio da expressão gênica, no qual um segmento de nucleotídios de uma das fitas do DNA serve de molde para a síntese de uma molécula de RNA complementar antiparalela no núcleo celular. Para ser realizada, a transcrição necessita de uma fita molde de DNA, de substratos para a confecção da molécula de RNA que será sintetizada e de um complexo proteico que inclui a presença de uma RNA polimerase dependente de DNA.

Do mesmo modo que a replicação do DNA, a transcrição também utiliza como base a *complementaridade entre nucleotídios*; contudo, algumas diferenças podem ser notadas entre os dois processos. Em primeiro lugar, enquanto na replicação os precursores são desoxirribonucleosídios trifosfatos, na transcrição o substrato utilizado para a confecção da molécula de RNA é representado pelos ribonucleosídios trifosfatos. A síntese da molécula de RNA também acontece no sentido 5′ → 3′, como na replicação do DNA. No entanto, enquanto na replicação as duas fitas do DNA são copiadas, na transcrição apenas uma das fitas do DNA (3′ → 5′) é geralmente utilizada pela maioria dos genes como molde para a síntese de uma molécula de RNA complementar com polaridade inversa. Nesse aspecto, a sequência da molécula de RNA sintetizada (5′ → 3′) será idêntica em sequência e polaridade à da fita do DNA que não foi utilizada como molde (5′ → 3′), exceto pelas uridinas no lugar das timidinas presentes no DNA. Por essa razão, a fita do DNA não molde é chamada de *fita senso*, e a fita molde (3′ → 5′) recebe a designação de *antissenso*. Nos bancos de dados disponíveis para consulta de sequências gênicas, a sequência do DNA é sempre representada pela fita senso (5′ → 3′).

Além das diferenças citadas, o processo de transcrição não precisa de um *primer* para ser iniciado, como ocorre na replicação, e, enquanto na replicação todos os nucleotídios da fita molde de DNA são copiados, na transcrição apenas partes da molécula de DNA, representando um gene individual ou mais raramente um segmento com poucos genes, são transcritas em dado momento. Isso ocorre porque nem todos os genes necessitam se expressar ao mesmo tempo em um tecido e em um período específico.

A transcrição pode ser dividida em três etapas principais: *início*, *alongamento* e *término*.

Embora o processo transcricional seja semelhante entre procariotos e eucariotos, existem peculiaridades entre os dois grupos de organismos; por esse motivo, eles serão abordados separadamente.

Transcrição em procariotos

O *início* da transcrição ocorre com o reconhecimento das sequências consenso que representam o promotor e a montagem do complexo de proteínas envolvido no processo transcricional. Em bactérias, as sequências-alvo do promotor são reconhecidas diretamente pela RNA polimerase; em eucariotos, proteínas de ligação ao DNA formam uma plataforma, por meio da qual a RNA polimerase pode se ligar ao DNA.

O promotor é o grande responsável por sinalizar a partir de qual nucleotídio a transcrição será iniciada, qual fita do DNA será utilizada como molde e em que direção a RNA polimerase irá se deslocar.

Na maioria das bactérias, o *promotor* é representado por duas sequências consenso anteriores ao ponto de início da transcrição: o *boxe -35* (5′ TTGACA 3′), localizado 35 nucleotídios antes do ponto +1, e o *boxe -10* (5′ TATAAT 3′), também conhecido como *Pribnow boxe*, localizado 10 nucleotídios antes do sítio de início da transcrição. Vários genes bacterianos funcionalmente relacionados encontram-se aglomerados, sob o controle de um único promotor e são transcritos como uma única e longa molécula de RNA.

Bactérias apresentam apenas um tipo de RNA polimerase, representada por uma grande estrutura multimérica formada por várias cadeias polipeptídicas individuais.

No cerne dessa enzima, responsável por sintetizar a molécula de RNA nascente, podem ser encontradas cinco cadeias distintas: duas cópias da subunidade alfa (α), uma subunidade

beta (β), uma subunidade beta primo (β') e uma subunidade ômega (ω). No entanto, outras subunidades podem se ligar e se desligar temporariamente ao cerne da enzima ao longo do processo. Uma dessas subunidades é o fator sigma (σ), que controla a ligação da RNA polimerase ao promotor e o início da transcrição, possibilitando que o processo seja iniciado no sítio correto. À associação do fator sigma ao cerne da RNA polimerase, é dado o nome de holoenzima, a qual se liga aos boxes -35 e -10, ocupando um espaço que vai desde o nucleotídio -50 até o nucleotídio +20, o que possibilita que o sítio ativo da RNA polimerase seja posicionado no ponto +1. Após o início do processo, o fator sigma (σ) se desliga do complexo.

Holoenzima
Forma de uma enzima multimérica, na qual todos os componentes polipeptídicos estão presentes

Não é necessária a presença de um *primer* para a incorporação do primeiro nucleotídio à fita de RNA nascente.

Ao se ligar ao promotor, a holoenzima desenrola a dupla hélice de DNA pelo rompimento das pontes de hidrogênio que unem as duas fitas em uma extensão que inclui o ponto de início da transcrição. Em seguida, por complementaridade, um nucleosídio trifosfato é pareado com o sítio +1 da fita de DNA molde. Após o início da transcrição, a RNA polimerase sofre mudança na conformação, o que a impede de permanecer ligada ao promotor. Tal mudança é essencial para a continuidade da transcrição, pois proporciona à enzima que se desloque ao longo da molécula de DNA.

A etapa de *alongamento* é caracterizada pela incorporação de ribonucleosídios monofosfatados, um a um, à hidroxila na extremidade 3′ da cadeia crescente de RNA por meio de ligações fosfodiéster, os quais são resultado da remoção de dois resíduos de pirofosfato do trifosfato do ribonucleosídio precursor apropriado. Assim, o primeiro ribonucleotídio incorporado na extremidade 5′ da fita de RNA nascente é diferente dos demais por apresentar um grupamento trifosfato 5′.

A RNA polimerase apresenta grande precisão na incorporação dos ribonucleosídios, e evidências demonstram que ela também tem capacidade de revisão. Ao incorporar um ribonucleosídio incorreto, a enzima é capaz de voltar e cortar os dois últimos nucleotídios integrados à molécula de RNA crescente, recomeçando o processo de transcrição a partir desse ponto. Com isso, o erro pode ser imediatamente reparado.

Capacidade de revisão
Habilidade das polimerases de remover e substituir nucleotídios incorporados incorretamente

À medida que a dupla hélice é desenrolada, a RNA polimerase se desloca. Todavia, apenas uma pequena região da dupla hélice sofre relaxamento por vez, formando uma bolha de transcrição. O deslocamento dessa bolha possibilita a movimentação da RNA polimerase ao longo da molécula de DNA e o alongamento do transcrito. Enquanto o DNA na frente da bolha de transcrição é desenrolado, a região pela qual a bolha acabou de passar sofre nova helicoidização pela ação de girases. Contudo, o deslocamento dessa bolha, como também ocorre na replicação do DNA, provoca uma tensão na molécula de DNA, que é aliviada pela atuação de topoisomerases.

Bolha de transcrição
Segmento da molécula de DNA que foi desenrolado para expor uma sequência unifilamentar, a qual será transcrita em RNA

O *término* da transcrição é caracterizado pelo reconhecimento da sequência finalizadora, seguida da interrupção da atividade da RNA polimerase e da subsequente separação da fita de RNA recém-sintetizada da molécula molde de DNA. Em geral, os finalizadores são encontrados antes do ponto em que o término da transcrição efetivamente ocorre, de modo que a transcrição termina somente após o finalizador também ter sido transcrito. Bactérias exibem dois tipos de finalizadores: *dependentes* e *independentes da proteína rô.*

Nos finalizadores independentes da proteína rô, a RNA polimerase reconhece por si mesma o fim da transcrição e os sítios de término. Esses finalizadores, denominados *terminadores intrínsecos,* apresentam sequências repetidas de nucleotídios na molécula de DNA que são invertidas e complementares, seguidas de 6 adeninas consecutivas após um intervalo de 7 a 9 pares de bases. Quando essas sequências invertidas geralmente ricas em CG são transcritas no RNA, elas se dobram imediatamente para parear entre si, formando uma estrutura secundária em formato de grampo. A presença dessa estrutura faz com que a RNA polimerase diminua a velocidade de transcrição ou faça uma pausa. Como a região rica em uracilas posteriores ao grampo tem pareamento de bases fraco (A-U), isso é suficiente para desestabilizar o híbrido RNA-DNA, o que propicia o término do processo transcricional.

Nos finalizadores dependentes, o fator rô se liga à molécula de RNA em um sítio anterior ao sítio de término e move-se ao longo do RNA em direção à extremidade 3′. A RNA polimerase encontra um grampo finalizador e diminui sua velocidade. Assim, a proteína rô a alcança e entra em ação na região híbrida DNA-RNA, atuando como uma helicase, capaz de separar a cadeia de RNA da fita molde de DNA na bolha de transcrição. Contudo, não se conhece ainda se essa atividade é suficiente para a liberação do transcrito ou se o fator rô interage com a RNA polimerase para que a liberação do RNA aconteça e a transcrição seja finalizada.

RNA policistrônico
Molécula de RNA transcrita de um grupo de genes contínuos capaz de promover a síntese de múltiplas proteínas

Após a dissociação do transcrito de RNA da molécula de DNA, ele pode exercer seu papel diretamente ou pode servir como mensageiro para a síntese de um polipeptídio por meio da tradução. Em bactérias, um conjunto de genes contínuos pode ser transcrito em uma única molécula de RNA, chamada de **RNA policistrônico**, porém esse evento é mais raro em eucariotos. Uma súmula das diferentes etapas que compõem o processo de transcrição em procariotos pode ser observada na Figura 2.3.

Transcrição em eucariotos

Nas células eucarióticas, a transcrição ocorre no núcleo e, em menor escala, nas mitocôndrias e nos cloroplastos. Embora as etapas de início, alongamento e término sejam semelhantes às que ocorrem em procariotos, o processo é bem mais complexo.

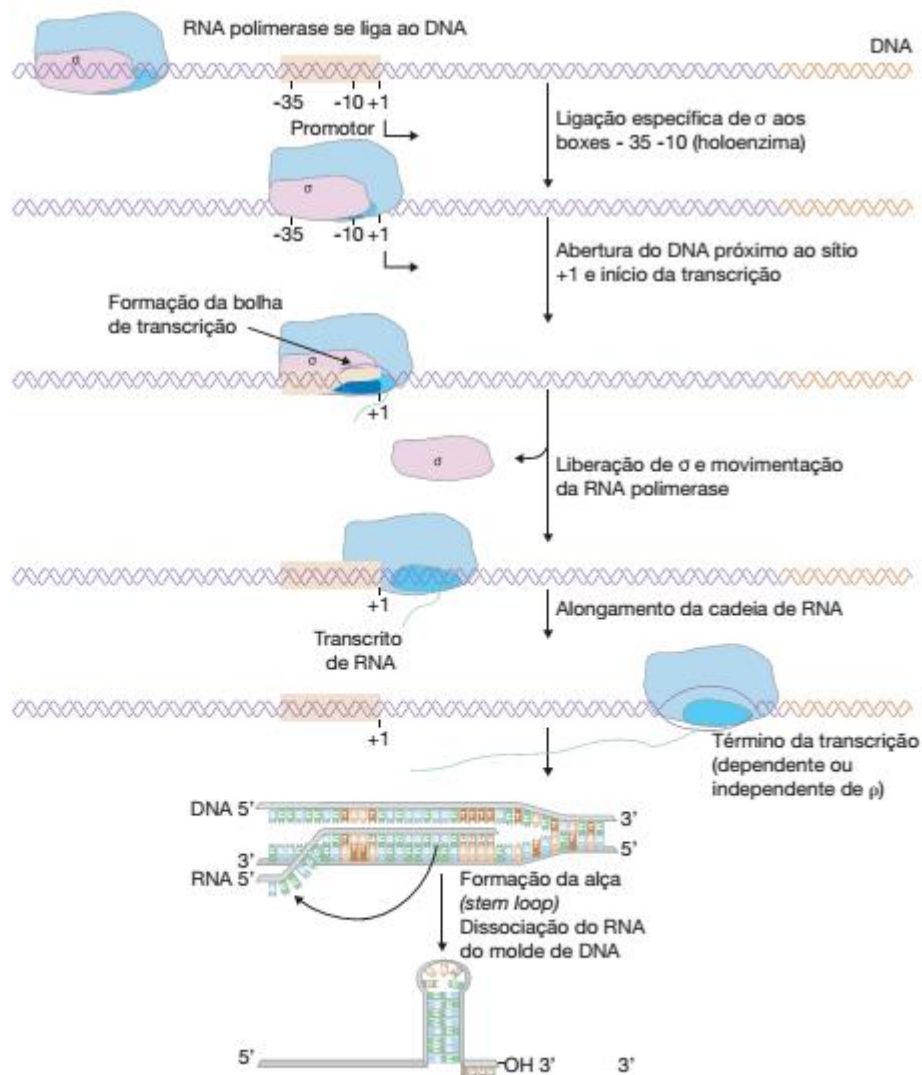



Figura 2.3 Visão geral das etapas de início, alongamento e término da transcrição em procariotos.

Na maioria dos eucariotos, 3 tipos de RNA polimerase (I, II e III) podem ser observados.

Cada RNA polimerase reconhece um tipo específico de promotor e identifica sinais de término específicos. Por esse motivo, as RNA polimerases apresentam especificidade para transcrever diferentes tipos de RNA. A RNA polimerase I é responsável por transcrever todos os rRNA, exceto o pequeno rRNA 5S; a RNA polimerase II transcreve mRNA, snoRNA, alguns snRNA e determinados miRNA. Já a RNA polimerase III transcreve tRNA, o pequeno rRNA 5S, certos miRNA e alguns snRNA. A localização das 3 RNA polimerases também está relacionada com os genes que elas transcrevem, de modo que a RNA polimerase I está localizada no nucléolo e as RNA polimerases II e III, no nucleoplasma. Em plantas, uma quarta RNA polimerase (RNA polimerase IV) também pode ser observada, sendo responsável por transcrever alguns siRNA. Assim como a RNA polimerase bacteriana, as RNA polimerases de eucariotos também são multiméricas, mas o número atuante de subunidades é bem maior.

Nucleoplasma
Parte do núcleo que exclui o nucléolo

Diferentemente do que ocorre em bactérias, nas quais as regiões promotoras são reconhecidas diretamente pela RNA polimerase, essas enzimas em eucariotos não têm autonomia para iniciar a transcrição por si mesmas.

Sendo assim, é necessário que proteínas se liguem ao DNA, construindo uma *plataforma* nas sequências promotoras, pela qual a RNA polimerase pode se ligar ao DNA. Essas proteínas reguladoras são denominadas *fatores de transcrição transatuantes*.

Os elementos de transcrição são chamados de transatuantes por serem produzidos por genes localizados a distância e terem de migrar para seus locais de atuação. Os elementos promotores aos quais esses fatores transatuantes se ligam à molécula de DNA são denominados *elementos cisatuantes*. Somente com a interação entre os fatores trans e cisatuantes, a RNA polimerase pode dar início à síntese de RNA.

A ligação dos fatores transcricionais ao DNA é, portanto, o que especifica quais genes serão expressos em dado momento e o que possibilita a síntese de transcritos de RNA a partir desses genes.

A ligação ao DNA não é a única função dos fatores transcricionais, que também podem atuar reconhecendo outros fatores, inclusive a própria RNA polimerase, podendo ser ainda incorporados em um complexo de iniciação na presença de várias outras proteínas. Tipos diferentes de promotores em eucariotos requerem a ligação de diferentes tipos de fatores transcricionais. Para as RNA polimerases I e III, os fatores transcricionais são relativamente simples, mas para a RNA polimerase II, eles formam um grande complexo proteico ao redor do ponto de início de transcrição, que, em conjunto com a RNA polimerase, é denominado aparato de transcrição basal.

Aparato de transcrição basal
Complexo de fatores de transcrição, RNA polimerase e demais proteínas montados no promotor que são capazes de iniciar a transcrição em níveis basais (mínimos)

Os promotores de eucariotos são bem mais complexos do que os de procariotos e consistem não só da sequência na qual a RNA polimerase se ligará ao DNA, como de outras sequências que atuam como reguladoras da expressão gênica, inibindo ou estimulando a transcrição. Os promotores das RNA polimerases I e II são majoritariamente localizados *antes do ponto de início da transcrição*, mas alguns promotores para a RNA polimerase III situam-se após esse ponto. Como cada RNA polimerase tem diferentes tipos de elementos promotores, abordaremos primeiramente os *relacionados com RNA polimerase II*, enzima utilizada pela maioria de nossos genes codificadores de proteínas.

Os promotores da RNA polimerase II podem ser classificados em diferentes tipos dispostos a seguir. Alguns desses elementos serão detalhados ao longo deste capítulo.

- *Promotores centrais*: são elementos que normalmente estão localizados antes do ponto de início da transcrição e direcionam o começo da transcrição. Na ausência de elementos reguladores adicionais, os promotores centrais possibilitam a expressão constitutiva do gene em níveis basais (mínimos). Esses promotores incluem o TATA Box ou Box -25, uma sequência consenso TATAAA, que se situa 25 pb antes do sítio de início da transcrição. O TATA Box tem o papel primordial de posicionar o ponto de início da transcrição
- *Promotores não centrais*: são elementos localizados anteriormente aos promotores centrais. Um exemplo é o CCAAT Box (sequência consenso GGCCAATCT), localizado 75 nucleotídios antes do ponto de início da transcrição

- *Reforçadores*: são elementos no DNA que servem para intensificar os níveis de transcrição quando ligados por proteínas reforçadoras ou ativadoras
- *Silenciadores*: são elementos no DNA que servem para diminuir os níveis de transcrição, quando ligados por proteínas silenciadoras
- *Elementos insuladores*: bloqueiam a ação de elementos que têm efeito reforçador ou silenciador sobre a transcrição
- *Elementos de resposta*: modulam a transcrição em resposta a um estímulo externo.

Para que a transcrição seja iniciada, é necessário que a *cromatina* esteja acessível à maquinaria transcricional.

Nucleosoomo
Unidade estrutural básica da cromotina composta pelo enovelamento do DNA a um cerne de oito proteínas histona (duas de cada H2A, H2B, H3 e H4)

Regulação epigenética
Forma de regulação gênica que envolve modificações no DNA estáveis e herdáveis que não incluem mudanças na sequência de bases

Como sabemos, nosso DNA está organizado em nucleossomos, formados pela interação, em grande parte, da molécula de DNA com proteínas histonas, o que faz com que a cromatina esteja condensada. Você deve estar pensando: como os fatores transcricionais conseguirão interagir e se ligar à molécula de DNA? Na realidade, a cromatina apresenta uma *estrutura dinâmica* que é regulada, em parte, por modificações químicas pós-traducionais nas caudas das histonas. Um exemplo dessas modificações inclui a acetilação de aminoácidos localizados nas extremidades das histonas. Esse processo desestabiliza a estrutura do nucleossomo, tornando o DNA ali presente mais acessível à maquinaria transcricional. Além disso, esse tipo de regulação epigenética, sobre a qual apresentaremos mais detalhes a seguir, pode ter a participação de proteínas de remodelagem da cromatina capazes de modificar o posicionamento individual de nucleossomos dos promotores e de outras regiões reguladoras da transcrição. Com isso, essas regiões se tornam livres para a ligação dos fatores transcricionais.

Isso demonstra que as histonas não representam apenas estruturas proteicas inertes, mas têm participação ativa na determinação dos genes que serão expressos.

Além disso, a metilação do DNA é um dos vários mecanismos reguladores que influenciam a atividade de um promotor. A adição de grupamentos metil na região promotora de um gene, preferencialmente em regiões ricas em dinucleotídios CG ligados covalentemente (ilhas CpG), é capaz de impedir a transcrição. Em virtude desse mecanismo, os grupamentos metil devem ser removidos dessas regiões para que o promotor possa se tornar ativo.

De acordo com o exposto, o início da transcrição pela RNA polimerase II em eucariotos se dá pela ligação de proteínas reguladoras ao DNA, em uma região próxima ao promotor, propiciando a descompactação da cromatina. O aparato transcricional pode, então, ser montado no promotor, sendo constituído pela RNA polimerase II e por uma gama de fatores transcricionais gerais, consistindo em 50 ou mais polipeptídios. Entre os fatores transcricionais estão TFIIA, TFIIB, TFIID, TFIIE, TFIIF E TFIIH (TF do inglês *transcription factor*; II de RNA polimerase II e A, B, D, E, F, H para designar cada fator individualmente). Os fatores gerais de transcrição têm função análoga ao fator sigma em bactérias. O contato inicial com as sequências promotoras é feito pelo TFIID, que se liga ao TATA Box na fita molde de DNA. O TFIID é formado por, no mínimo, 9 polipeptídios, entre os quais está a proteína de ligação à TATA (TBP). A TBP se liga ao sulco menor da molécula de DNA, unindo-se a ela como se fosse uma sela, deselicoidizando-a parcialmente. Em seguida, os outros fatores ligam-se ao complexo e à RNA polimerase, posicionando a enzima no sítio de início da transcrição.

Após a montagem da RNA polimerase e dos fatores transcricionais no promotor, tanto o DNA quanto a polimerase sofrem mudanças conformacionais, resultando em separação da dupla fita de DNA em uma pequena extensão que circunda o sítio de início da transcrição. A transformação desse complexo em uma forma aberta possibilita que a fita molde de DNA fique acessível e a síntese de RNA possa ser iniciada.

Em eucariotos, a primeira base do RNA a ser transcrita é uma adenina (A), flanqueada em ambos os lados por sequências de pirimidinas.

A região é chamada de *iniciadora* (Inr), sendo representada pela sequência consenso Py_2CAPy_5, localizada no sítio de início de transcrição (Py equivale a uma pirimidina).

As RNA polimerase I e III iniciam a transcrição de maneira similar, mas bem mais simples que aquela utilizada pela RNA polimerase II.

Os *promotores das RNA polimerases I e III* são bem diferentes dos usados pela RNA polimerase II. Eles estão majoritariamente localizados *antes* do ponto de início da transcrição, mas alguns promotores para a RNA polimerase III situam-se após esse ponto, sendo chamados de *regiões de controle interno*. Quando o aparato de transcrição é formado, a RNA polimerase III vai para o começo do gene e inicia a transcrição.

Após a RNA polimerase ter sintetizado cerca de 30 pb da cadeia de RNA crescente, a enzima deixa o promotor e inicia o processo de *alongamento*. Alguns fatores transcricionais permanecem no promotor à espera de uma nova RNA polimerase para reinício de uma nova rodada de transcrição. Da mesma maneira como acontece em bactérias, a RNA polimerase se move ao longo do DNA pela formação da bolha de transcrição, acrescentando novos ribonucleosídios monofosfatados à extremidade 3′ da cadeia crescente de RNA.

Para o *término* da transcrição, as RNA polimerases I, II e III também utilizam mecanismos diferentes. O término da transcrição de genes transcritos pela RNA polimerase II envolve a exonuclease Rat1.

Em geral, a RNA polimerase II transcreve centenas e até mesmo milhares de nucleotídios após a sequência necessária para a síntese do RNA.

Enquanto a RNA polimerase ainda está transcrevendo, ocorre uma clivagem endonucleolítica na molécula de RNA em uma sequência consenso, separando a molécula de RNA em dois segmentos, o que codificará a proteína e um outro, cuja extremidade 5′ servirá para ligação da exonuclease Rat1. Após a ligação da Rat1, ela se deslocará em direção à RNA polimerase que continua a transcrever, degradando o segundo RNA na direção 5′ → 3′. Quando a Rat1 alcança a polimerase, a transcrição é finalizada.

Em contraste com a RNA polimerase II, as RNA polimerase I e III fazem uso de diferentes sinais de término. A RNA polimerase I utiliza um fator de término semelhante ao fator rô utilizado na transcrição bacteriana. No entanto, enquanto o fator rô se liga à cadeia de RNA que está sendo sintetizada, o fator de término para a RNA polimerase I se liga a um segmento da molécula de DNA localizado após o ponto de término. Já a RNA polimerase III utiliza como elemento finalizador uma região que produz uma série consecutiva de uracilas na molécula de RNA, como aquela utilizada no término da transcrição bacteriana na ausência da proteína rô. Contudo, para a RNA polimerase III não há a formação de uma estrutura de grampo prévia a esse ponto. Na Figura 2.4, todas as etapas que compõem a transcrição em eucariotos estão resumidamente ilustradas.

■ Tipos e processamento dos RNA

O RNA é uma molécula com estrutura e função versáteis. Nesse sentido, como abordamos anteriormente, diferentes tipos de RNA podem ser produzidos pela transcrição (Figura 2.5).

O mRNA, o rRNA e o tRNA são os principais tipos de RNA celulares em meio a vários outros tipos de moléculas menores de RNA.

Curiosamente, embora o mRNA codifique todas as proteínas, ele representa apenas 0,5% a 1% de todo o RNA celular gerado pelo genoma humano. A sequência de nucleotídios de uma molécula de RNA transcrita pode conter regiões que são complementares entre si, possibilitando o pareamento intramolecular entre eles pela formação de pontes de hidrogênio. A combinação de regiões pareadas e não pareadas possibilita que a sequência unifilamentar de RNA adquira, portanto, estruturas secundárias complexas, como a formação de grampos e hastes, que são importantes para sua estabilidade e funcionalidade (Figura 2.6).

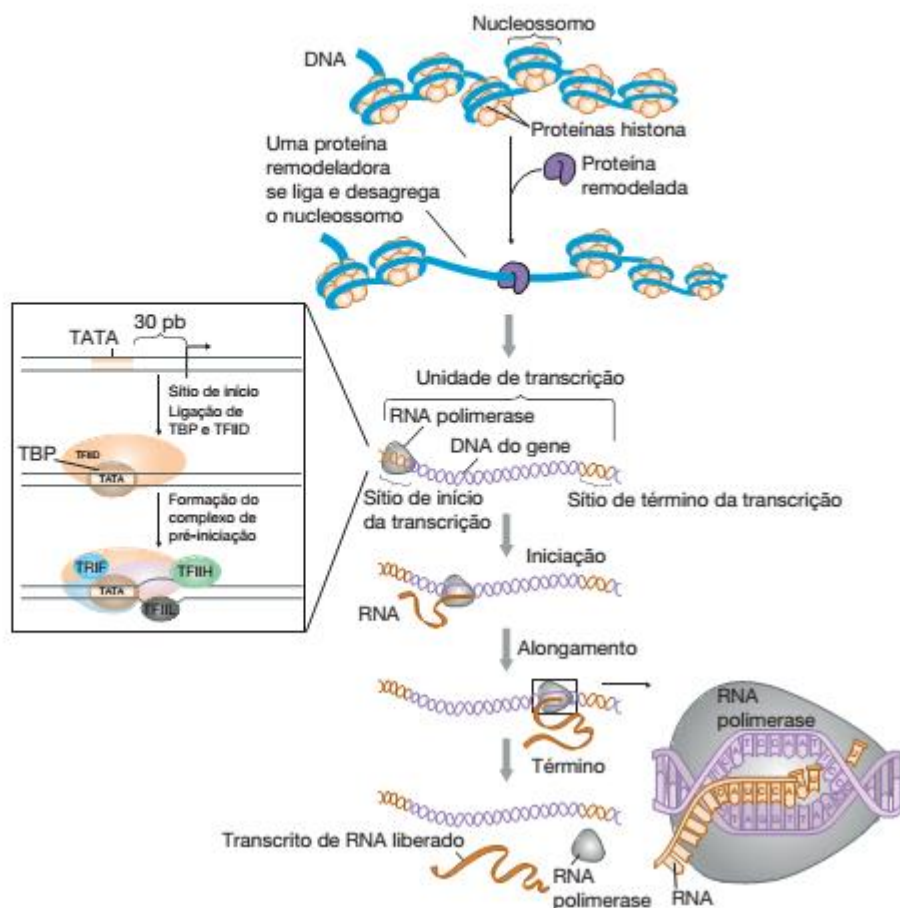



Figura 2.4 Visão geral das etapas de início, alongamento e término da transcrição em eucariotos.

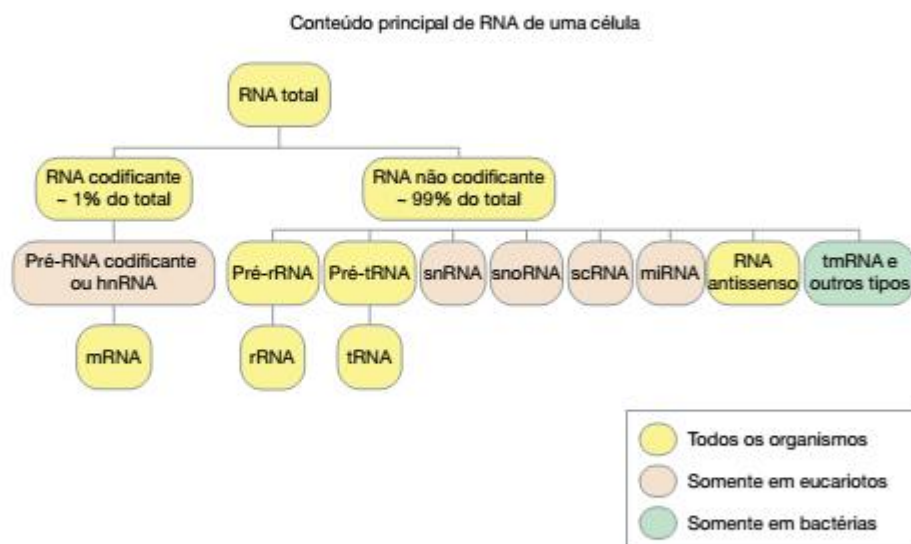



Figura 2.5 Principais tipos de RNA que podem ser produzidos durante a transcrição.

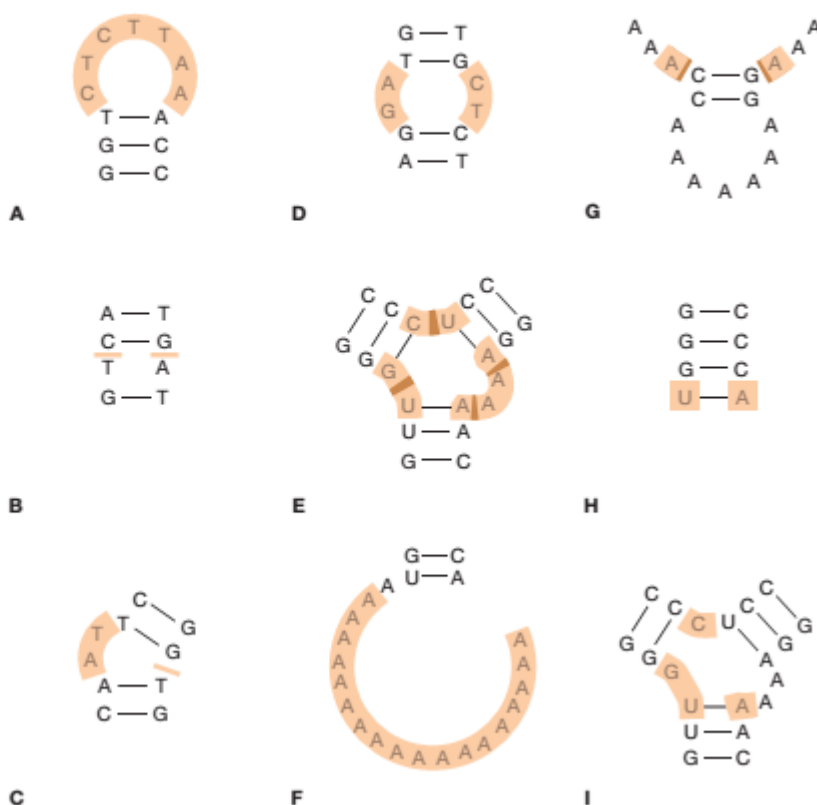



Figura 2.6 Estruturas secundárias em forma de haste e alça originadas na molécula de RNA.

O mRNA e seu processamento

Transcrito primário
Molécula de RNA eucariótica que é modificada após a transcrição até gerar uma molécula de RNA madura

Para genes, cujo produto gênico final é uma proteína, o transcrito inicial de RNA, chamado de transcrito primário, *pré-mRNA* ou *RNA heterogêneo nuclear* (hnRNA), deve sofrer diferentes modificações que possibilitam sua migração do núcleo eucariótico, onde ocorreu a transcrição, até os ribossomos citoplasmáticos, onde ocorrerá a tradução. Já em procariotos, a transcrição e a tradução ocorrem em um espaço comum e de maneira acoplada, de modo que enquanto a extremidade 3′ da molécula de RNA ainda está sendo transcrita, a outra extremidade (5′) já começa a ser traduzida, não havendo tempo hábil para modificações do transcrito. Sendo assim, o mRNA procariótico que será traduzido é uma cópia fiel do gene.

As modificações do pré-mRNA eucariótico são representadas por: (a) modificações químicas em ambas as extremidades do transcrito (capeamento da extremidade 5′ e poliadenilação da extremidade 3′); (b) remoção dos íntrons da sequência codificante com recomposição subsequente dos éxons a eles adjacentes (encadeamento do RNA) e, (c) em alguns casos, a edição do mRNA.

O *capeamento do transcrito primário* ocorre logo após o início da transcrição e representa a adição de um nucleotídio modificado, uma *7-metilguanosina*, na *extremidade 5′* da molécula de RNA.

Capeamento 5′
Modificação na extremidade 5′ de um mRNA eucariótico pela adição de um nucleotídio modificado quimicamente, importante para o trajeto do núcleo ao citoplasma e para a ligação dos ribossomos durante a tradução

Como abordado, o primeiro ribonucleosídio adicionado à molécula de RNA crescente no início da transcrição é o único que apresenta três grupamentos fosfato ligados ao carbono 5′. No capeamento 5′, um desses grupamentos fosfato terminais é removido por uma fosfatase RNA terminal e um nucleotídio guanina com seu grupamento fosfato é adicionado à extremidade 5′ do pré-mRNA, por uma ligação trifosfato 5′-5′ não usual. Essa reação é catalisada por uma enzima guanilil transferase, como demonstrado na Figura 2.7. Em seguida, um grupamento metil (CH_3) é adicionado à posição 7 da base do nucleotídio guanina terminal, transformando-a em 7-metilguanosina trifosfato pela ação de uma guanina metiltransferase. Além disso, pode ocorrer ainda a adição de radicais metil por outras metiltransferases na posição 2′ do açúcar ribose no segundo e terceiro nucleotídios subsequentes.

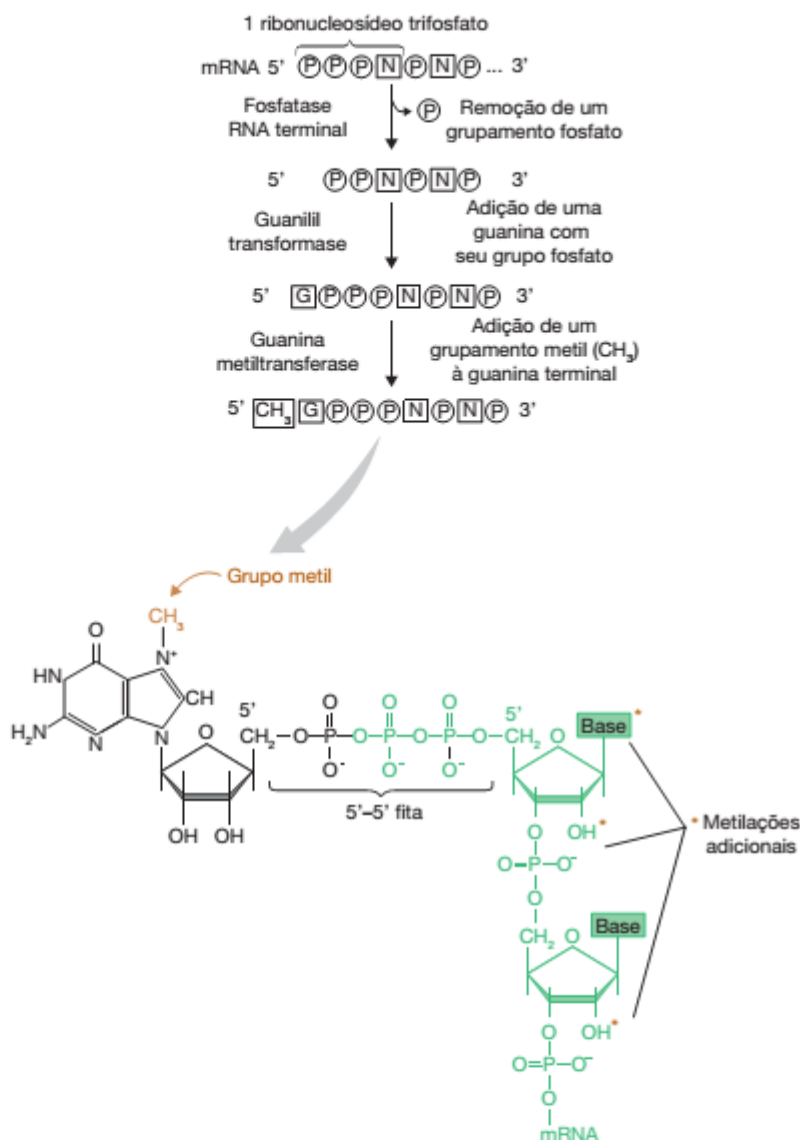



Figura 2.7 Capeamento na extremidade 5′ do transcrito de pré-mRNA eucariótico. Os asteriscos indicam as posições nas quais metilações adicionais podem ocorrer.

As enzimas necessárias para o capeamento se associam especificamente ao complexo de transcrição da RNA polimerase II, utilizada pelos genes codificadores de proteínas durante a transcrição.

Como as RNA polimerases I e III, responsáveis pela síntese de rRNA, tRNA e alguns snRNA, não se associam a essas enzimas, esses três tipos de RNA não são modificados por capeamento.

O capeamento tem diferentes funções que envolvem: (a) facilitar a exportação do transcrito do núcleo para o citoplasma; (b) facilitar a recomposição do transcrito; (c) proteger o transcrito do ataque de endonucleases com atividade 5′→ 3′ e (d) promover o encaixe do ribossomo ao mRNA durante a tradução.

Em seguida ao capeamento na extremidade 5′, outra modificação química ocorre na extremidade 3′ e é essencial para a estabilidade do transcrito: a poliadenilação. Essa modificação consiste na adição após a transcrição de 50 a 250 adeninas à extremidade 3′ (poliadenilação 3′), gerando uma estrutura chamada de *cauda poli(A)*. Vários genes eucarióticos transcritos pela RNA polimerase II transcrevem vários nucleotídios após o final da região que será codificada. Nesses casos, a sequência adicional é clivada tomando-se por base o reconhecimento de sequências consenso que estão antes e depois do sítio de clivagem. A sequência AAUAAA em geral se localiza 11 a 30 nucleotídios antes do corte e sinaliza a posição perante à qual o transcrito deve ser clivado para a adição da cauda poli (A). Após o sítio de clivagem, existe

Poliadenilação 3′
Adição de resíduos de adenina subsequentes à extremidade 3′ de um mRNA que tem grande importância em garantir estabilidade ao mRNA

uma sequência rica em uracilas. As funções da poliadenilação incluem estabilizar o mRNA no citoplasma, já que tamanhos pequenos de caudas poli(A) estão associados à degradação da molécula de RNA, e facilitar o reconhecimento do mRNA pelos ribossomos. Há indícios de que as modificações químicas finais nas extremidades 5′ e 3′ são finalizadas antes da remoção dos íntrons, mas isso pode não ser a regra para todos os mRNA. Alguns genes também podem utilizar, de modo alternativo, múltiplos sítios de clivagem 3′ para a adição da cauda poli(A), o que é chamado de poliadenilação alternativa. A utilização de um sítio de clivagem alternativo pode ou não levar a diferentes isoformas proteicas, dependendo da localização anterior ou posterior do sítio em relação ao códon de término da tradução.

Poliadenilação alternativa
Utilização diferencial de sítios distintos de clivagem na extremidade 3′ de um transcrito de mRNA para a adição da cauda poli(A)

Ainda no núcleo celular, um terceiro tipo de modificação acontece no transcrito. Essa modificação consiste na recomposição (*splicing*) do pré-mRNA, caracterizada pela remoção dos íntrons e união (encadeamento) dos éxons adjacentes. A remoção efetiva de um íntron depende do reconhecimento de 3 sequências consenso principais: duas sequências localizadas nos limites entre um íntron e um éxon (uma no início e outra no fim) e 1 *sítio de ramificação*.

Recomposição (*splicing*) do RNA
Etapa do processamento do mRNA da maioria dos genes de eucariotos, a qual consiste na excisão dos íntrons e no encadeamento dos éxons do transcrito primário com a participação do spliceossomo

As sequências conservadas localizadas nos limites íntron/éxon geralmente incluem uma sequência GU no início do íntron (5′) e uma sequência AG no final do íntron (3′). A terceira sequência consenso essencial para a recomposição do pré-mRNA inclui uma adenina em uma posição que antecede de 18 a 40 nucleotídios o sítio de corte na extremidade 3′ do íntron. Mutações pontuais nessas 3 sequências podem impedir que a remoção dos íntrons ocorra de maneira adequada, levando à permanência, no mRNA maduro, de sequências não codificantes, as quais passam a ser lidas como codificantes durante a tradução (Capítulo 3). Atualmente, são conhecidas várias doenças causadas por mutações em sequências-consenso intrônicas, utilizadas para remoção dos íntrons.

Nos últimos anos, tem sido observado que os éxons também contêm sequências que funcionam como sítios de ligação para proteínas que atuam como *moduladoras* do processo de *splicing*. Entre essas sequências, estão os *sítios exônicos acentuadores do splicing* (*e*xonic *splicing e*nhancers, ou ESEs), que são motivos de 6 a 8 nucleotídios, capazes de definir os éxons para a maquinaria do *splicing*. Mutações que inativem ESEs podem resultar, portanto, em *exclusão de éxons* (*exon skipping*), podendo estar associadas a doenças (Capítulo 3).

Spliceossomo
Complexo de pequenos RNA nucleares que participam da recomposição do pré-mRNA após a transcrição de genes eucarióticos

A recomposição do pré-mRNA ocorre dentro de um grande complexo molecular de RNA e proteínas chamado de spliceossomo. Essa estrutura reúne 5 moléculas de *pequenos RNA nucleares* (snRNA) chamados de U1, U2, U4, U5 e U6, os quais se associam a várias proteínas para formar *pequenas partículas de ribonucleoproteínas nucleares* (*snRNP*). Além das diferentes snRNP, o spliceossomo também apresenta proteínas que não se associam aos snRNA.

snRNA
Pequenos RNA nucleares que fazem parte da recomposição do pré-mRNA

Como ilustrado na Figura 2.8, o mecanismo de recomposição do mRNA envolve 3 etapas. Em primeiro lugar, ocorre clivagem na sequência de corte GU localizada no início do íntron (5′). Em seguida, a extremidade 5′ livre se liga ao sítio de ramificação, localizado dentro do íntron, formando uma estrutura em forma de laço. A formação dessa estrutura requer a hidrólise de ATP e é mediada por um ataque nucleofílico que envolve o grupamento hidroxila 2′ ligado à adenina conservada no sítio de ramificação e à guanina presente na sequência GU conservada no início do íntron, resultando em uma ligação covalente entre esses 2 nucleotídios. Por último, a sequência consenso no final do íntron (3′) é clivada e o íntron em laço é liberado, fazendo com que os dois éxons adjacentes possam se unir covalentemente. A união dos éxons, em geral, ocorre na mesma ordem que eles se apresentavam no transcrito primário. O íntron liberado é linearizado após a quebra de ligação no sítio de ramificação e, posteriormente, degradado por enzimas nucleares. Curiosamente, pelo fato de o gene da distrofina humana ser muito longo e demorar muito para ser transcrito, o RNA gerado sofre recomposição dos éxons antes mesmo de a transcrição estar completamente finalizada.

Além dos íntrons nucleares de pré-mRNA, cujo mecanismo de recomposição foi descrito, são reconhecidos mais 3 tipos principais de íntrons: *íntrons do grupo I*, *íntrons do grupo II* e *íntrons de RNA transportador*. Em bactérias, os íntrons do grupo I estão presentes em genes de rRNA, mRNA e tRNA, ao passo que, em eucariotos inferiores, eles podem ser encontrados em genomas de mitocôndrias e de cloroplastos e em genes nucleares de rRNA. Em plantas superiores, os íntrons do grupo I estão restritos a alguns genes de tRNA e mRNA em cloroplastos

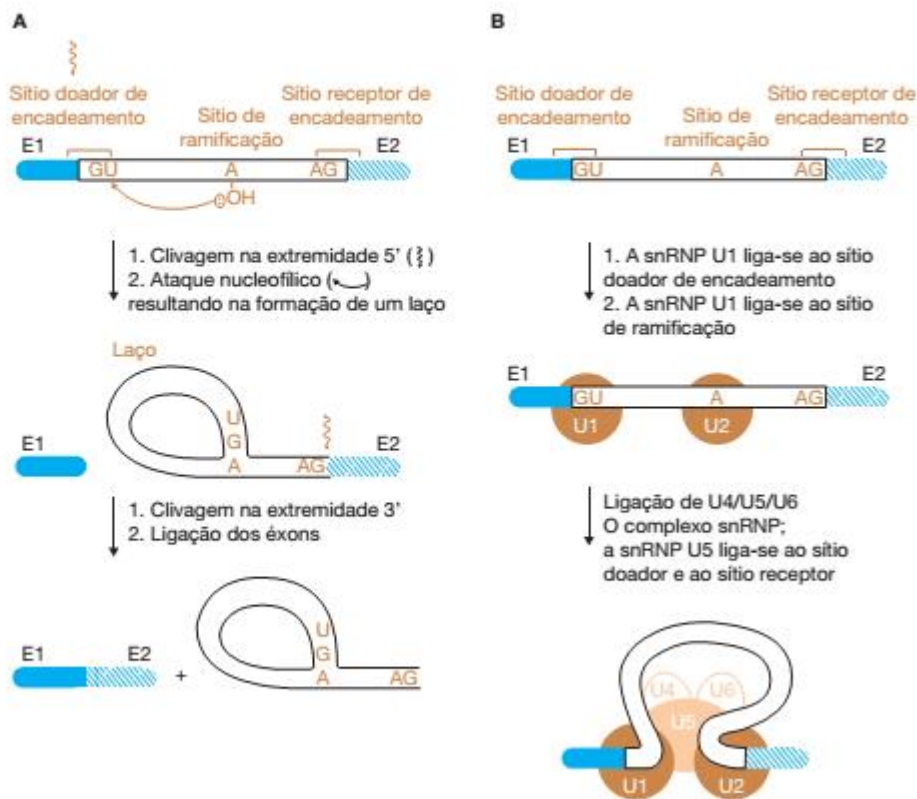



Figura 2.8 Processo de recomposição dos éxons no pré-mRNA eucariótico.

e mitocôndrias. Estes têm a habilidade de catalisar sua *autorremoção*, sem o envolvimento de uma atividade enzimática proteica. Embora apresentem tamanhos diversificados, os íntrons do grupo I têm estrutura secundária comum formada por 9 alças e hastes.

Os íntrons do grupo II também podem ser autoprocessados e são observados em alguns genes de rRNA, tRNA e mRNA, em organelas de fungos, plantas e protistas, como também

G·F Genética em Foco

Múltiplas facetas do genoma humano

Além do encadeamento alternativo do RNA, o genoma humano pode utilizar ainda outras artimanhas para codificar vários produtos gênicos com base em um número reduzido de genes. Em nosso genoma, encontramos em média 1 gene a cada 40 kb a 45 kb de DNA. Tendo em vista que os genes humanos têm em média de 10 kb a 15 kb de comprimento, podemos considerar que nossos genes estão separados por regiões intergênicas de cerca de 30 kb. No entanto, em casos raros, podemos encontrar genes no genoma humano que se sobrepõem parcialmente em sequência. Alguns desses genes são também transcritos com base em diferentes fitas da molécula de DNA. A região de classe III do complexo HLA, localizada na região cromossômica 6p21.3, é um exemplo de local com alta densidade gênica com sobreposição de genes. No genoma mitocondrial, também podemos encontrar exemplos de genes sobrepostos. Além destes, também existem situações nas quais genes pequenos estão localizados dentro de genes maiores. Exemplos dessa disposição são representados pelos genes que codificam RNA nucleolar (snoRNA) e o gene da neurofibromatose do tipo I (NF1), no qual um dos íntrons (íntron 27) contém 3 pequenos genes, transcritos com base na fita oposta à que é utilizada para a transcrição do NF1.

em genes bacterianos de mRNA. No entanto, de modo contrário ao que acontece com os íntrons de grupo I, a excisão de íntrons do grupo II ocorre na ausência de GTP e envolve a formação de um laço, semelhante ao mecanismo envolvido na excisão dos íntrons nucleares de pré-mRNA. Por último, os íntrons de RNA transportador, como o nome já nos possibilita saber, são aqueles localizados nos genes de tRNA. Os íntrons nos pré-tRNA de bactérias são autorremovidos, enquanto que em *archae* e eucariotos esse subgrupo de íntrons é removido por endonucleases de recomposição. As principais etapas que compõem o processamento do pré-mRNA estão resumidas na Figura 2.9.

Recomposição ou encadeamento alternativo (*splicing* alternativo) União combinatória diferenciada de conjuntos de éxons durante a recomposição do pré-mRNA que favorece a formação de mais de um produto com base em um único gene

Alguns genes podem sofrer ainda recomposição ou encadeamento alternativo (*splicing* alternativo), mecanismo pelo qual ocorre uma união combinatória diferencial dos éxons, originando diferentes sequências de aminoácidos com base em um mesmo transcrito de pré-mRNA. Isso faz com que um mesmo gene possa produzir distintas isoformas de proteínas, que podem também ter pequenas diferenças funcionais e especificidade de tecido. Estima-se que cerca de 60% de todos os genes humanos sofram recomposição alternativa do transcrito, ocasionando grande diversidade de proteínas. Este é considerado um dos mecanismos que poderia explicar o número reduzido de genes na espécie humana (cerca de 30.000) em comparação com outras espécies menos evoluídas (Capítulo 1).

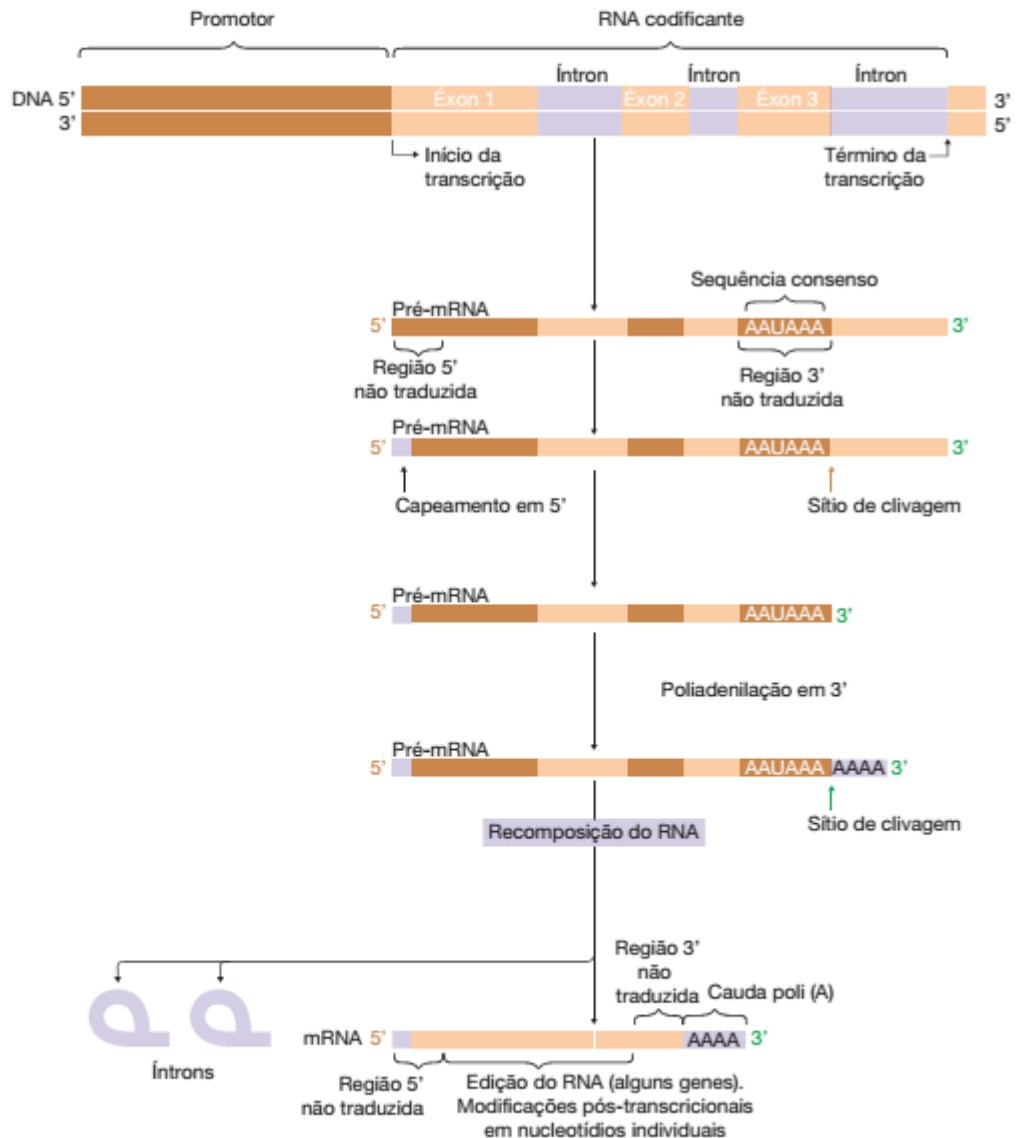



Figura 2.9 Súmula das principais etapas de processamento do pré-mRNA eucariótico.

Edição do RNA
Processo natural raro representado pela modificação da sequência de um transcrito de RNA, seja pela adição de grupamentos químicos, seja pela alteração de nucleotídios originalmente presentes

Somados ao capeamento na extremidade 5′, à poliadenilação na extremidade 3′ e à recomposição, alguns transcritos são submetidos ainda à edição do RNA. Embora se utilize o conceito de que o transcrito de RNA é uma cópia fiel da fita molde de DNA, esse mecanismo possibilita haver modificações pós-transcricionais em nucleotídios individuais localizados no mRNA. Isso faz com que a proteína codificada tenha uma sequência de aminoácidos diferente daquela informação genética contida inicialmente no gene. Em alguns casos, ocorre conversão de bases por modificações químicas. Por exemplo, uma desaminação de uma citosina a converte em uma uracila, podendo alterar o aminoácido que seria codificado originalmente pelo códon anterior. Em outros casos, a edição do RNA conta com a participação de uma *molécula guia de RNA* (gRNA) capaz de parear com o mRNA pré-editado e possibilitar que haja adição, deleção ou alteração de nucleotídios na região do pareamento, de acordo com a sequência presente no gRNA.

O rRNA e seu processamento

As moléculas de rRNA são componentes essenciais dos ribossomos, organelas complexas abundantes na maioria das células que atuam como as grandes fábricas de proteínas em nosso organismo. Além de conter cerca de 80% de todo o RNA celular, os ribossomos também são constituídos por mais de 50 proteínas diferentes. O tamanho dos ribossomos é estimado por sua velocidade de sedimentação após centrifugação, medida em unidades Svedberg (S). Em procariotos, o ribossomo tem um coeficiente de sedimentação de 70S, enquanto em eucariotos esse coeficiente é de 80S.

Cada ribossomo é formado por duas subunidades funcionais (*subunidades maior e menor*), distintas em tamanho. Em *E. coli*, essas subunidades são de 50S e 30S. A subunidade maior contém 2 moléculas de rRNA (23S e 5S) e a subunidade menor apresenta apenas 1 molécula de rRNA de 16S. Os ribossomos eucarióticos têm subunidades de 60S e 40S. A subunidade maior contém 3 moléculas de rRNA (28S, 5,8S e 5S) e a menor tem apenas 1 molécula de rRNA de 18S. Cada ribossomo tem 1 cópia de cada 1 das diferentes moléculas de rRNA, sendo, portanto, 3 para o ribossomo de procariotos e 4 para aquele de eucariotos.

Você deve estar se perguntando: por que os valores das velocidades de sedimentação não são aditivos? Na realidade, a velocidade de sedimentação é afetada pela massa e pela estrutura tridimensional dos ribossomos, de modo que os valores de sedimentação referentes a cada subunidade, se somados, não representam o valor de sedimentação real observado para o ribossomo como um todo.

Os genes que codificam rRNA costumam estar presentes em múltiplas cópias e tanto as moléculas de rRNA bacterianas, como as eucarióticas sofrem processamento.

Em procariotos, os genes de 16S, 23S e 5S são transcritos juntos por uma mesma unidade de transcrição, gerando um pré-rRNA de 30S (Figura 2.10). Esse pré-rRNA é modificado quimicamente, dando início a bases não usuais em pontos específicos. A principal modificação química do pré-rRNA é a adição de grupamentos metil em sítios específicos. Após a modificação, o pré-rRNA é, então, clivado para gerar os rRNA maduros de 16S, 23S, 5S.

Em eucariotos, os genes de rRNA 5,8S, 18S e 28S estão dispostos consecutivamente nas regiões organizadoras de nucléolos dos cromossomos (Capítulo 4), enquanto os genes de rRNA 5S encontram-se distribuídos ao longo de vários cromossomos. Dessa forma, os genes de rRNA 28S, 18S e 5,8S de eucariotos também são transcritos juntos com base em uma única unidade de transcrição, gerando um transcrito de rRNA de 45S que é posteriormente processado. Todavia, o gene do rRNA de 5S está localizado em um outro sítio, sendo transcrito de maneira independente dos demais rRNA.

snoRNA
Pequenos RNA nucleolares que participam do processamento dos rRNA eucarióticos

O processamento dos rRNA eucarióticos ocorre nos nucléolos de modo semelhante ao processo de procariotos. No entanto, as modificações químicas dos pré-rRNA envolvem a participação de *pequenas moléculas de RNA nucleolares* (snoRNA), que se associam a proteínas

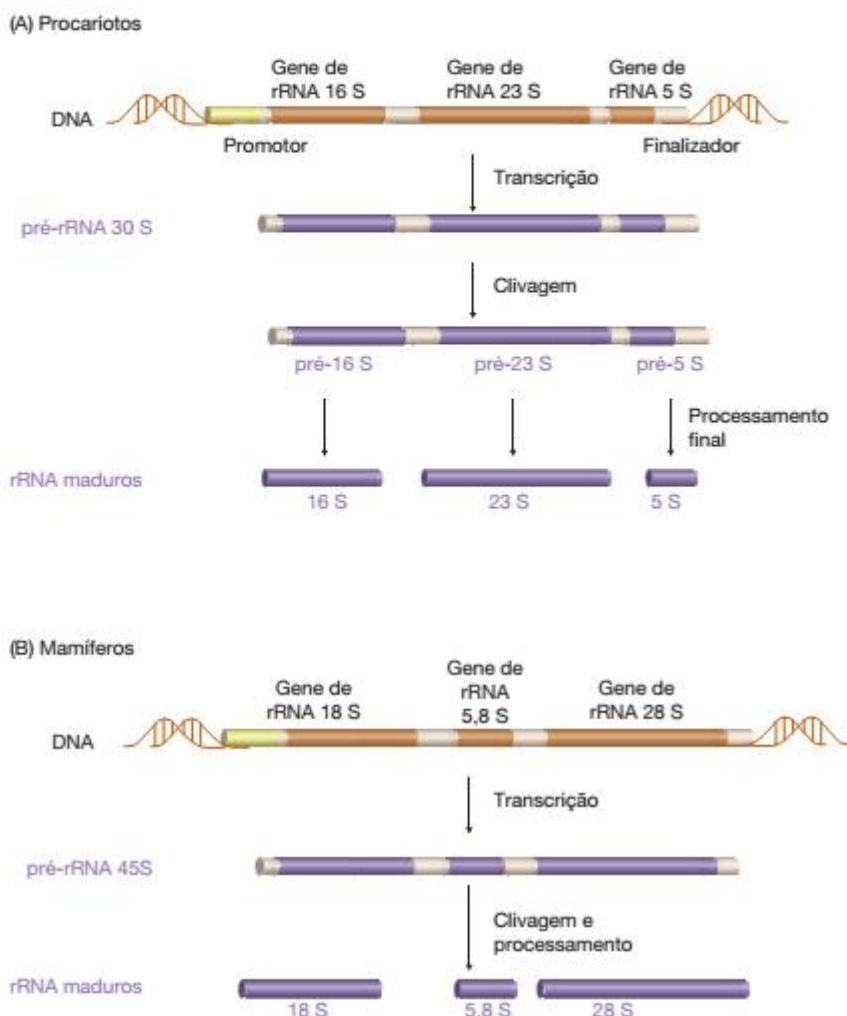



Figura 2.10 Processamento do rRNA. Em procariotos (A) e em mamíferos (B). Em mamíferos, os rRNA 5S são transcritos a partir de uma unidade independente.

formando *pequenas partículas de ribonucleoproteínas*. Por apresentarem complementariedade à sequência do pré-rRNA, os snoRNA pareiam com sítios-alvo, indicando os locais do pré-rRNA nos quais devem acontecer modificações químicas. Os snoRNA não devem ser confundidos com os pequenos RNA nucleares (snRNA) utilizados no processamento do mRNA.

O tRNA e seu processamento

O tRNA é uma molécula adaptadora, que funciona como um elo entre as informações contidas em um transcrito de mRNA e a proteína que será por ele codificada. São as moléculas de tRNA que leem o conteúdo genético presente na sequência de nucleotídios do mRNA e a convertem em uma linguagem de aminoácidos. Essa leitura da informação genética contida na molécula de mRNA é feita através dos códons, que correspondem a sequências de 3 nucleotídios. Cada códon é reconhecido pelo tRNA correspondente por complementariedade de bases, sendo decodificado em um aminoácido.

Códon
Sequência de 3 nucleotídios adjacentes capazes de especificar um aminoácido durante a tradução

Cada tRNA liga-se especificamente a um tipo de aminoácido, o leva até o ribossomo e o adiciona à cadeia polipeptídica crescente. A denominação do tRNA dependerá, portanto, do aminoácido que ele carrega. Um tRNA que transporta uma arginina será denominado $tRNA^{Arg}$. Como existem mais tipos diferentes de tRNA em uma célula (cerca de 30 a 50) do que os 20 aminoácidos existentes, alguns aminoácidos são transportados por mais de um tipo de tRNA. Esses tRNA diferentes que levam um mesmo aminoácido são chamados de *tRNA isoaceptores*.

As moléculas de tRNA são relativamente pequenas (74 e 95 nucleotídios), similares em tamanho e conservadas, assumindo a mesma configuração secundária em *forma de trevo* após sua transcrição, como pode ser observado na Figura 2.11. Essa estrutura diferenciada com 4 braços principais formados por pareamento intramolecular entre nucleotídios complementares é essencial para que o tRNA possa exercer sua função de molécula adaptadora.

Os braços da estrutura em trevo são denominados de braço aceptor, braço TψC, braço anticódon e braço DHU.

O *braço aceptor*, que inclui as extremidades 5′ e 3′ da molécula de tRNA, e representa o local onde o aminoácido específico se liga. Na extremidade 3′ desse braço existe uma sequência de nucleotídios 5′ CCA 3′ que está presente em todos os tRNA, não tendo, portanto, relação com a especificidade do aminoácido que será ligado no braço aceptor. Essa sequência 3′ terminal é adicionada após a transcrição da molécula de tRNA por uma enzima tRNA nucleotidil transferase. O *braço TψC* recebe esse nome por apresentar uma alça em sua extremidade contendo uma sequência de timina, (T), uma pirimidina modificada chamada de pseudouridina (ψ) e uma citosina (C). Por sua vez, o *braço anticódon*, localizado na posição oposta ao braço aceptor, é aquele que apresenta sequência complementar ao códon que está sendo lido. A estrutura do braço anticódon apresenta em sua alça 3 nucleotídios que constituem o anticódon. O anticódon varia para cada tRNA e irá parear com o códon correspondente no mRNA, garantindo que cada aminoácido seja incorporado na ordem correta na cadeia polipeptídica nascente. Por esse motivo, podemos dizer que o braço anticódon desempenha papel essencial na decodificação da informação genética presente na molécula de mRNA. O *braço DHU* é situado na posição oposta ao braço TψC e também tem uma pirimidina modificada em sua alça, a di-hidroxiuracila (DHU). Além desses 4 braços principais, pode ser encontrada ainda uma *quinta alça* variável em tamanho, contendo poucos nucleotídios, chamada de *braço extra*. Embora sempre observemos nos livros didáticos o tRNA sendo representado por um trevo, estudos cristalográficos evidenciam que a estrutura terciária do tRNA é semelhante a um L, formado pela dobradura da estrutura em forma de trevo sobre si mesma. Nessa conformação tridimensional, os braços aceptor e anticódon estão localizados em posições opostas da molécula.

Anticódon
Sequência de 3 bases em uma molécula de tRNA que pareia com um códon específico no mRNA

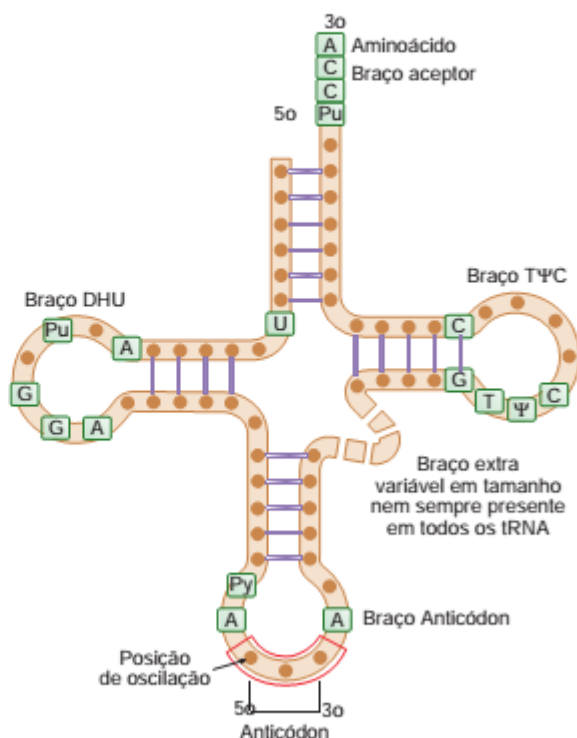



Figura 2.11 Estrutura do tRNA, evidenciando os braços aceptor, TΨC, anticódon e DHU.

Como acontece com o mRNA, o tRNA também é transcrito como um pré-tRNA, que é posteriormente processado, tanto em bactérias como em eucariotos, até originar a forma madura de tRNA, pronta para exercer sua função. Em bactérias, vários tRNA podem ser transcritos juntos em uma única unidade de transcrição como uma grande molécula, que é subsequentemente clivada em pontos específicos para dar origem a várias moléculas maduras independentes de tRNA. A clivagem das extremidades 5′ e 3′ dos trevos ocorre por ação das ribonucleases *RNase P* e *RNase D*, respectivamente, seguida da inserção de uma sequência CCA 3´ terminal no braço aceptor por ação da tRNA nucleotidil transferase (Figura 2.12). Em bactérias, alguns genes que produzem tRNA estão em cópia única, enquanto outros são encontrados em múltiplas cópias. Em virtude da demanda da síntese de tRNA em eucariotos, nesses organismos geralmente cada gene de tRNA apresenta muitas cópias.

Alguns genes de tRNA em *archae* e em eucariotos apresentam íntrons que são removidos durante o processamento dos pré-tRNA. Contudo, em comparação com os íntrons observados nos pré-mRNA, os presentes em pré-tRNA costumam ser menores e não apresentam sequências consenso nos limites íntron/éxon. Assim, o mecanismo de encadeamento desses pré-tRNA ocorre também de maneira distinta dos transcritos que codificarão proteínas. Além disso, a modificação química de alguns nucleotídios presentes na molécula de tRNA após sua transcrição é um evento comum e parece estar associado à estabilidade da estrutura do tRNA, garantindo que nenhum aminoácido seja incorporado de maneira incorreta na cadeia polipeptídica. Dessa forma, após o processamento, modificações químicas no transcrito de t-RNA podem ocorrer com frequência bem superior àquela observada para o rRNA. No entanto, diferente do que vimos na edição do mRNA, na qual as sequências de nucleotídios originais presentes em um gene podem ser editadas originando produtos proteicos diferentes, as modificações químicas na sequência de nucleotídios dos tRNA e rRNA após sua transcrição influenciam somente a estrutura de suas moléculas. As modificações mais comuns no tRNA incluem:

- a adição de grupamentos químicos (p. ex., adição de grupamentos metil à guanosina gerando 7-metilguanosina)
- desaminação de bases (p. ex., guanosina, gerando inosina)
- substituição do enxofre
- conversão de uma ligação dupla por uma simples (saturação de dupla ligação)
- isomerização de bases.

Outros tipos de RNA

miRNA
Pequeno RNA de interferência produzido pela clivagem de um RNA geralmente unifilamentar, cuja estrutura se dobra em forma de um grampo bifilamentar. Os miRNA se conjugam a complexos de proteínas para regular a expressão gênica em nível pós-transcricional de mRNAs-alvo

siRNA
Pequeno RNA de interferência unifilamentar produzido pela clivagem de um RNA bifilamentar que pode ter origem exógena. Atua na regulação pós-transcricional da regulação gênica ao parear com moléculas de RNA-alvo complementares

Além dos 3 principais tipos de RNA (mRNA, rRNA e tRNA), cujos processamentos foram detalhados até o momento, vimos que moléculas de RNA nucleares (snRNA) e nucleolares (snoRNA) pequenas têm papel essencial no processamento do mRNA e do rRNA, respectivamente. Além dessas, na última década, uma série de outras moléculas de RNA não codificadores foi descrita. Embora pequenos, estes RNA atuam como reguladores vitais da expressão gênica e, em alguns casos, como determinantes hereditários da variação epigenética. A descoberta desse grupo crescente de RNA não só traz nova luz sobre a regulação gênica pós-transcricional, como também possibilita explicar, em parte, como surgem algumas doenças que fogem ao padrão clássico de herança mendeliano.

Uma classe de pequenos RNA que tem recebido atenção especial nos últimos anos é formada pelos microRNA (**miRNA**) e os pequenos RNA de interferência (**siRNA**) que, embora se originem de modo diferente, têm várias similaridades em seus mecanismos de atuação. Ambos os tipos de RNA estão interligados pelo processo de *interferência por RNA*, utilizado pelas células eucariotas para proteger o organismo da invasão de genes exógenos, como os de vírus e transposons, bem como para modular negativamente a expressão gênica.

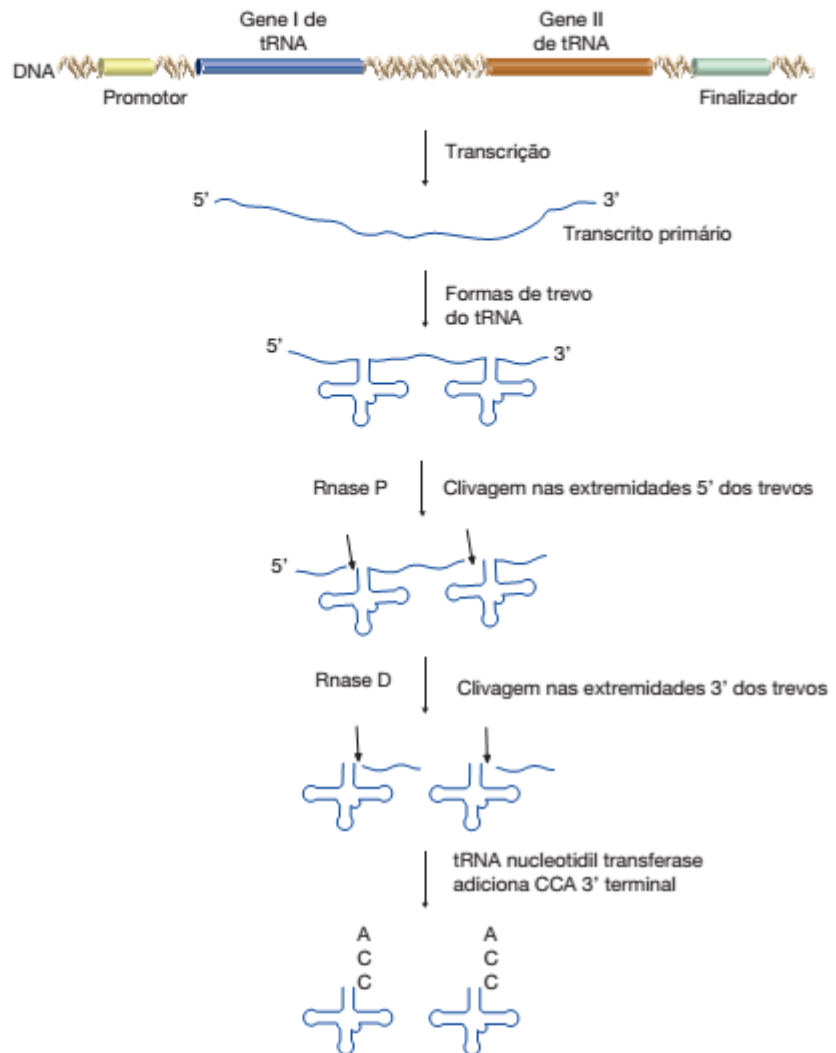



Figura 2.12 Transcrição e processamento das moléculas de tRNA, enfocando a atuação das enzimas Rnase P, Rnase D e tRNA nucleotidil transferase.

A interferência por RNA ocorre inicialmente por uma molécula bifilamentar de RNA, que pode ser formada pela transcrição de um gene que apresenta sequências intramoleculares complementares, as quais possibilitam a dobradura do transcrito unifilamentar; pela transcrição simultânea de duas moléculas distintas de RNA que se complementam e pareiam formando um duplex de RNA ou pela infecção por um vírus com RNA. Essas moléculas de RNA em fita dupla são processadas para formar moléculas de fita simples maduras de miRNA ou siRNA.

Os miRNA consistem em moléculas endógenas de RNA de fita simples com 19 a 25 nucleotídios em tamanho, os quais modulam a expressão gênica majoritariamente por meio do silenciamento pós-transcricional de mRNAs-alvo.

Os miRNA são expressos amplamente, mas alguns têm especificidade de tecido e de estágio do desenvolvimento. Em geral, um único miRNA pode regular vários mRNA-alvo. Os miRNA estão envolvidos em vários processos biológicos, como desenvolvimento embrionário, ciclo celular, diferenciação hematopoética, apoptose, prevenção da inflamação, espermatogênese e desenvolvimento neuronal. Em humanos, mais de 2042 miRNA foram descritos, embora esse número seja crescente. Os *loci* de miRNA geralmente são encontrados agrupados e aqueles miRNA localizados em um mesmo módulo podem estar

funcionalmente relacionados. Os genes de miRNA podem ser expressos a partir de íntrons de unidades de transcrição codificadoras e não codificadoras de proteínas, de regiões intergênicas, de éxons de mRNA ou a partir de sequências que codificam unicamente miRNA.

Os miRNA são inicialmente transcritos no núcleo, preferencialmente pela RNA polimerase II, como *longos miRNA primários* (*pri-miRNA*) que variam em tamanho de centenas de bases até 10 kb. Esses transcritos primários se dobram em formato de um grampo de cabelo e são processados em *miRNA precursores* (*pré-miRNA*) por um complexo catalítico que envolve a participação de uma RNAse III denominada Drosha e de uma proteína de ligação ao RNA de fita dupla (DGCR8). Os pré-miRNA são, então, exportados para o citoplasma pela exportina 5, onde são processados em moléculas de *miRNA maduros*. A transformação dos pré-miRNA em miRNA maduros é regida pela enzima Dicer, outro tipo de RNase III que cliva a alça terminal do grampo. Com a separação do duplex de RNA, o filamento de RNA com menor estabilidade, denominado *miRNA** é quase sempre degradado, enquanto o outro filamento, chamado de miRNA, é acoplado em uma partícula ribonuclear para formar um complexo de silenciamento induzido por RNA (RISC). Em seguida, o miRNA maduro guia o complexo RISC até sítios complementares presentes em geral na extremidade 3′ não traduzida dos transcritos de mRNA-alvo. A sequência do miRNA que pareia precisamente com o mRNA tem 7 nucleotídios e é denominada sequência *seed*. Ao encontrar seus mRNA-alvo, dois processos podem acontecer: repressão da tradução ou degradação do mRNA complementar, ambos levando ao silenciamento do gene cujo mRNA foi alvo de um miRNA. A escolha por cada um desses mecanismos dependerá, em parte, do grau de complementaridade entre o miRNA e os sítios de ligação no mRNA-alvo e dos componentes proteicos do complexo RISC (Figura 2.13).

Complexo de silenciamento induzido por RNA
Conjugação de proteínas à uma molécula de miRNA ou siRNA que possibilita a regulação pós-transcricional negativa de moléculas de mRNA-alvo

Os pequenos RNA de interferência (siRNA) também têm cerca de 22 nucleotídios de tamanho, mas costumam surgir de clivagens de transcritos de mRNA, de RNA de transpósons e de vírus com RNA. Assim, enquanto os miRNA são produzidos unicamente por longos transcritos primários unifilamentares de RNA que se dobram em uma estrutura

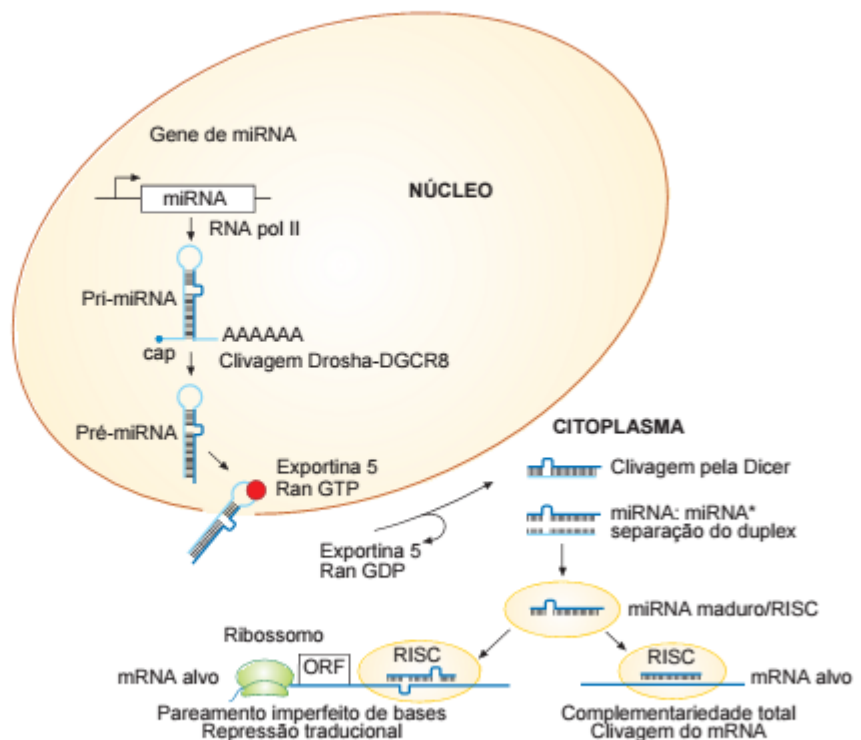



Figura 2.13 Biogênese resumida dos microRNA; transcrição no núcleo como pri-miRNA, processamento em pré-miRNA, exportação para o citoplasma, processamento pela Dicer e acoplamento ao complexo RISC e por fim, repressão dos mRNA-alvo.

bifilamentar, os siRNA surgem também por longos duplex de RNA, que podem ter origem exógena.

O processamento dos siRNA também envolve a participação da Dicer e o acoplamento do siRNA maduro ao complexo RISC. No entanto, enquanto um mesmo miRNA pode regular vários mRNA-alvo produzidos por genes também distintos daqueles que deram origem a esse miRNA, os siRNA em geral silenciam os próprios genes a partir dos quais foram transcritos. Portanto, podemos dizer que os siRNA são mais alvo-específicos. O processo de silenciamento também parece ter algumas distinções, com alguns siRNA ocasionando degradação do RNA-alvo, enquanto outros afetam a transcrição por meio de alterações na estrutura da cromatina. É válido ressaltar que essas diferenças entre miRNA e siRNA mudam a cada momento, com novas evidências experimentais.

Além dos miRNA e siRNA, outras classes de RNA incluem: os pequenos RNA citoplasmáticos (scRNA), que são encontrados no retículo endoplasmático rugoso do citosol, associados a proteínas envolvidas na seleção e no transporte de outras proteínas, e os RNA de interação Piwi (piRNA), os quais interagem com as proteínas Piwi, participando na regulação da espermatogênese, sendo encontrados nos testículos de mamíferos.

Tradução

A tradução é o processo pelo qual uma molécula de mRNA é decodificada nos ribossomos para especificar a síntese de um polímero formado pela união de várias unidades monoméricas de aminoácidos ligados em série (um polipeptídio). Os polipeptídios são o segundo maior componente dos organismos vivos, ficando atrás apenas das moléculas de água em termos de massa total. Grande parte de nosso genoma não é codificada, representando sequências não codificadoras dos próprios genes, sequências repetitivas não funcionais de DNA, pseudogenes e fragmentos de genes que permaneceram no genoma ao longo do processo evolutivo.

Pseudogenes
Cópias não funcionais de genes que perderam sua capacidade de gerar moléculas de RNA e/ou proteínas ao longo da evolução

Os três principais tipos de RNA celulares produzidos durante a transcrição (mRNA, rRNA e tRNA) atuam em conjunto na tradução para produzir um polipeptídio.

Na tradução, a molécula de mRNA é a responsável por direcionar a síntese de uma cadeia polipeptídica, cuja sequência de aminoácidos é determinada pela sequência de nucleotídios presente no mRNA, que, por conseguinte, é derivada da sequência de nucleotídios contida no seu molde de DNA. Por outro lado, as moléculas de tRNA são as responsáveis por reconhecer a sequência de nucleotídios presente no mRNA e correlacioná-la com aminoácidos correspondentes. Por último, todo o processo acontece nos ribossomos, que são formados por cerca de 3 a 5 moléculas de rRNA e 50 a 90 proteínas diferentes, funcionando como grandes fábricas moleculares de proteínas. Os ribossomos são capazes de posicionar corretamente os tRNA com o mRNA, catalisando as ligações entre os aminoácidos que são adicionados ao polipeptídio nascente.

As regras que regem o processo de leitura dos nucleotídios na molécula de mRNA e a decodificação dos nucleotídios em aminoácidos compõem o que chamamos de *código genético*.

Código genético
Conjunto de regras que governa a tradução de uma molécula de mRNA em uma proteína

A primeira característica do código genético é que ele é lido em *trincas* consecutivas, ou seja, cada sequência de três nucleotídios, chamada de *códon*, será responsável por especificar um aminoácido. Conforme citado anteriormente, são reconhecidos 20 tipos de aminoácidos comuns, dos quais derivam mais algumas formas modificadas. A estrutura comum de um aminoácido é representada por um átomo central de carbono ligado a um grupo amino, um átomo de hidrogênio, um grupo carboxila e um grupo radical R que será diferente para cada um dos aminoácidos. Os grupos R são altamente variados, podendo ser hidrofóbicos, hidrofílicos, acídicos ou básicos, o que provoca uma infinita versatilidade estrutural e funcional para as proteínas. Os aminoácidos se unem uns aos outros por *ligações peptídicas*

Cadeia polipeptídica
Longas sequências de aminoácidos unidos por ligações peptídicas

para formar uma cadeia polipeptídica, que têm extremidades com polaridades distintas: uma contendo um grupamento amino livre (NH_3^+) e outra com um grupamento carboxila livre (COO-). Uma ou mais cadeias polipeptídicas formarão uma proteína que poderá ter uma função regulatória, estrutural ou enzimática.

A leitura do código genético em trincas somente foi esclarecida na década de 1960 por Francis Crick e seus colaboradores. Considerando que cada posição no mRNA pode representar uma de quatro bases possíveis (A, U, C ou G), se cada nucleotídio especificasse um aminoácido, teríamos apenas 4^1=4 combinações possíveis, não atendendo à codificação dos 20 aminoácidos já conhecidos até então. Caso o código genético fosse lido em duplas de nucleotídios, da mesma forma teríamos apenas 4.4=4^2 ou 16 combinações possíveis, um número também inferior ao número total de aminoácidos conhecidos. No entanto, se cada aminoácido fosse especificado por uma trinca de nucleotídios, o número de combinações possíveis (4.4.4=4^3 ou 64 combinações) seria mais do que suficiente para codificar os 20 aminoácidos conhecidos. A partir dessa hipótese, estudos de mutagênese em bacteriófagos foram conduzidos por Crick e seus colegas comprovando que cada sequência de 3 nucleotídios especifica um aminoácido.

Como existe um excesso de códons (64) em relação ao número total de aminoácidos (20), dizemos que o código genético é *degenerado* ou *redundante*, ou seja, mais de uma trinca de nucleotídios especifica um mesmo aminoácido.

Dessa forma, todos os aminoácidos, com exceção da metionina e do triptofano, são especificados por mais de um códon. A degeneração do código genético envolve com mais frequência a terceira base do códon. Por esse motivo, mutações na terceira base de um códon são menos propensas a causarem um fenótipo anormal (Capítulo 3). Códons que especificam um mesmo aminoácido são chamados de *códons sinônimos*. Como pode ser observado na Figura 2.14, dos 64 códons existentes, 3 não especificam nenhum aminoácido, funcionando apenas como

Segunda base	Primeira base			
	U	C	A	G
U	UUU, UUC } Phe; UUA, UUG } Leu	CUU, CUC, CUA, CUG } Leu	AUU, AUC, AUA, AUG } Ile; Met	GUU, GUC, GUA, GUG } Val
C	UCU, UCC, UCA, UCG } Ser	CCU, CCC, CCA, CCG } Pro	ACU, ACC, ACA, ACG } Thr	GCU, GCC, GCA, GCG } Ala
A	UAU, UAC } Tyr; UAA Terminação; UAG Terminação	CAU, CAC } His; CAA, CAG } Gln	AAU, AAC } Asn; AAA, AAG } Lys	GAU, GAC } Asp; GAA, GAG } Glu
G	UGU, UGC } Cys; UGA Terminação; UGG Trp	CGU, CGC, CGA, CGG } Arg	AGU, AGC } Ser; AGA, AGG } Arg	GGU, GGC, GGA, GGG } Gly

Símbolos de uma letra para aminoácidos

A Alanina	C Cisteína	D Ácido aspártico	E Ácido glutâmico	F Fenilalanina
G Glicina	H Histidina	I Isoleucina	K Lisina	L Leucina
M Metionina	N Asparagina	P Prolina	Q Glutamina	R Arginina
S Serina	T Treonina	V Valina	W Triptofano	Y Tirosina
X Significa um códon de terminação				

Figura 2.14 Código genético.

Códon de término
Códons que sinalizam para a finalização da tradução de um polipeptídio, sendo representados pelas sequências de base UAA, UAG e UGA

Fase aberta de leitura ou ORF
Sequência de DNA de um gene que é traduzida em uma proteína

Aminoacilação
Processo prévio ao início da tradução que liga um aminoácido a seu tRNA correspondente

sinais de reconhecimento para o fim da tradução, sendo denominados **códons de término** (UAA, UAG ou UGA).

> Durante a tradução, todas as trincas de nucleotídios são lidas consecutivamente sem que haja sobreposição na leitura dos códons, ou seja, um nucleotídio participa apenas de um códon.

O segmento de um gene codificante de uma proteína que é traduzido em uma proteína é chamado de fase aberta de leitura ou **ORF** (do inglês, *open reading frame*). Para qualquer mRNA, existem 3 modos em potencial de lermos o mRNA, ou seja, 3 fases abertas de leitura podem ser utilizadas, como podemos observar na Figura 2.15.

> A matriz de leitura correta a ser empregada durante o processo de tradução é ditada pelo reconhecimento de um *códon de início* da tradução. Em geral, esse códon é representado pela sequência AUG e é responsável por especificar uma metionina. Sendo assim, a ORF correta no DNA começa com o códon de início ATG e termina com um dos 3 códons de término (TAA, TAG ou TGA).

Além de o código genético ser lido em trincas e ser degenerado, podemos dizer ainda que ele é quase que totalmente *universal*. Isso quer dizer que quase todos os códons especificam o mesmo aminoácido nos diferentes organismos. Em algumas das exceções, códons de terminação passam a ser lidos como codificantes enquanto em outros códons codificantes passam a codificar aminoácidos diferentes daqueles em geral especificados na maioria dos organismos. Exceções ao código universal foram descritas no genoma mitocondrial, em bactérias e no genoma nuclear de protozoários. Nas mitocôndrias humanas e de outros mamíferos, por exemplo, AUA especifica metionina em vez de isoleucina, UGA é um códon de triptofano em vez de término de cadeia e AGG e AGA representam códons de término e não de arginina.

Agora que já abordamos as características do código genético, detalharemos o processo de tradução em si.

Antes de a tradução ser iniciada, um aminoácido deve se ligar a seu tRNA correspondente. Essa ligação, chamada de **aminoacilação**, acontece pela ligação de alta energia (covalente) entre o grupo carboxila (COO-) do aminoácido e a extremidade 3′-OH do braço aceptor do tRNA.

Essa ligação ocorre mais precisamente no carbono 2′ ou no carbono 3′ da adenina presente na sequência CCA.

Como a sequência CCA está presente na estrutura de todos os tRNA, você deve estar se perguntando: de que maneira podemos garantir a especificidade da ligação do tRNA ao aminoácido correto? Essa acurácia é gerenciada pela atuação de enzimas denominadas *aminoacil-tRNA sintetases*. Essas enzimas reconhecem o aminoácido correto de acordo com o grupamento R, o tamanho e a carga, além de identificarem o tRNA apropriado pela sequência de nucleotídios nele contida. A maioria dos organismos tem 20 tipos diferentes de aminoacil tRNA sintetase, um para cada tipo de aminoácido.

A aminoacilação envolve gasto de energia e acontece em dois passos subsequentes. Primeiramente, o aminoácido reage com ATP, produzindo um aminoacil-AMP e pirofosfato. Em seguida, o AMP é liberado e o aminoácido é transferido para o tRNA apropriado. Considerando que, em média, a maioria dos organismos tem entre 30 e 50 tRNA diferentes, isso quer

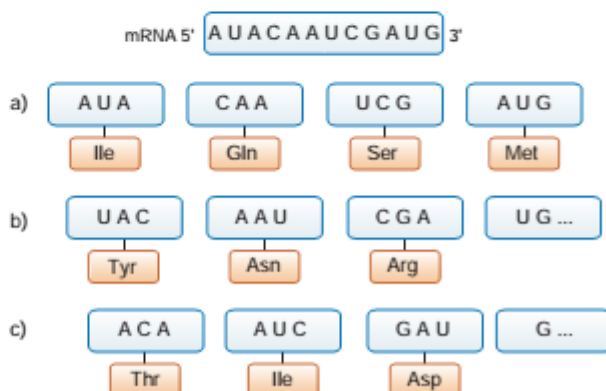



Figura 2.15 Exemplo hipotético das possíveis fases abertas de leitura de um mesmo transcrito de mRNA.

dizer que tRNA que se ligam a um mesmo aminoácido (tRNA isoaceptores) utilizam a mesma aminoacil-tRNA transferase.

O processo de aminoacilação é muito preciso em virtude da *capacidade de revisão* das aminoacil-tRNA sintetases, que podem remover aminoácidos incorporados incorretamente no tRNA. Sendo assim, a aminoacilação representa o primeiro nível de especificidade de ligação ao aminoácido correto, mas não é o único. Enquanto o braço aceptor do tRNA é responsável por transportar o aminoácido correspondente ao tRNA, o braço anticódon do tRNA é o que reconhecerá o códon presente no mRNA e se ligará a ele por complementaridade durante a tradução. Essa ligação representa o segundo nível de especificidade da tradução, possibilitando que o aminoácido correto seja incorporado na molécula polipeptídica nascente.

Considerando que existe ainda um excesso de códons no mRNA em relação ao número de anticódons presentes no tRNA, isso significa dizer que mais de um códon pode parear com um mesmo braço anticódon do tRNA. Nesse aspecto, existe uma *flexibilidade de pareamento* de bases entre o terceiro nucleotídio do códon do mRNA e o primeiro nucleotídio do anticódon do tRNA. Esse ponto de flexibilidade recebe o nome de posição de oscilação (Figura 2.11), e nela pode ocorrer um pareamento de bases incomum. Os exemplos mais comumente encontrados são o pareamento de bases G-U e o pareamento da inosina, uma guanosina desaminada, com C, A ou U. A posição de oscilação possibilita, dessa maneira, que uma célula possa sintetizar polipeptídios sem a necessidade de ter um grande número de tRNA.

Posição de oscilação
Pareamento entre o terceiro nucleotídio do códon do mRNA e o primeiro nucleotídio do anticodon do tRNA, no qual podem existir interações incomuns entre bases

Após o aminoácido correto ter sido ligado ao tRNA, podemos, então, dar início à tradução. A tradução envolve a ligação do ribossomo na extremidade 5′ do mRNA e seu deslocamento em direção à extremidade 3′, à medida que traduz os códons, um a um, durante a sua movimentação.

O processo de tradução em si envolve 3 etapas principais: (a) *início*, no qual ocorre montagem do aparato de tradução no ribossomo; (b) *alongamento*, etapa na qual os aminoácidos são acrescidos um a um à molécula polipeptídica nascente; e (c) *término*, no qual a síntese é finalizada com a identificação do códon de término e os componentes do aparato da tradução são desmembrados. Embora a estrutura dos ribossomos seja semelhante em eucariotos e bactérias, esses estágios apresentam diferenças importantes e, portanto, serão abordados de maneira independente.

Tradução em procariotos

Nas células bacterianas, a transcrição e a tradução ocorrem em um mesmo espaço e de maneira acoplada, já que, enquanto a extremidade 3′ de um mRNA está ainda sendo transcrita, os ribossomos já se ligam à extremidade 5′ do transcrito para iniciar a tradução. Como vimos, os ribossomos são formados por duas subunidades funcionais que, em bactérias, são representadas pelas subunidades 30S (subunidade menor) e 50S (subunidade maior). Em um ribossomo que não está sendo traduzido, as subunidades encontram-se unidas entre si. Sendo assim, o primeiro estágio da tradução (*início*), no qual o aparato traducional é montado, é evidenciado pela ligação do fator 3 de iniciação (IF3) à subunidade menor. Essa ligação impede que a subunidade maior se acople à subunidade menor, possibilitando que a molécula de mRNA que será traduzida se ligue à subunidade menor. No entanto, o ribossomo não pode se ligar a qualquer parte do mRNA. A ligação deve ocorrer precisamente em uma região específica do mRNA próxima ao primeiro códon a ser traduzido, o códon de início AUG. Em bactérias, essa região é representada pela sequência de Shine-Dalgarno, uma sequência conservada (5′-AGGAGGU-3′) localizada na região 5′UTR, antes do primeiro nucleotídio que será traduzido. A sequência Shine-Dalgarno apresenta complementaridade de bases com a sequência de nucleotídios localizada na extremidade 3′ do rRNA 16S da subunidade ribossômica menor. O pareamento entre a sequência Shine-Dalgarno e o rRNA 16S é imprescindível, portanto, para o posicionamento da subunidade menor no ponto correto.

Sequência de Shine-Dalgarno
Sequência consenso encontrada na região 5′ não traduzida do mRNA de bactérias que contém o sítio de ligação do ribossomo

Após a ligação da subunidade menor à sequência Shine-Dalgarno, a subunidade 30S migra até encontrar o códon AUG, localizado geralmente cerca de 10 nucleotídios após o sítio de ligação. Em seguida, o primeiro tRNA aminoacilado pareia com o códon de início AUG, pela

participação do fator 2 de iniciação (IF2), assim como de GTP, uma pequena molécula que libera energia ao ser convertida em difosfato de guanosina (GDP). Nessa etapa, o fator 1 de iniciação (IF1) também se liga à subunidade menor e parece ter a função de prevenir a ligação prematura de outros tRNA.

O primeiro tRNA que pareia com o códon de início carrega uma metionina que foi *modificada*, após a aminoacilação, pela substituição de um dos átomos de hidrogênio do grupamento amino por um grupamento formil (-COH).

Essa modificação resulta em uma *N-formilmetionina (fMet)* e impede que o grupamento amino da N-formilmetionina possa formar uma ligação peptídica, fazendo com que a polimerização do polipeptídio a ser traduzido ocorra obrigatoriamente na direção amino → carboxila. Nessa etapa, está montado o complexo de iniciação 30S formado pelo mRNA, pela subunidade ribossômica menor, pelo $tRNA^{fMet}$, pela molécula de GTP e pelos diferentes fatores de iniciação.

Depois que o aparato de início da tradução foi montado, os fatores de iniciação se dissociam do complexo e a molécula de GTP é hidrolisada em GDP. Com isso, a segunda etapa da tradução pode ser iniciada, o *alongamento*, caracterizado pela adição dos aminoácidos à molécula polipeptídica nascente. A partir de então, a subunidade maior (50S) se acopla ao aparato e são produzidos três sítios ribossômicos separados, nos quais os tRNA podem se ligar: o *sítio peptidil ou P*, o *sítio aminoacil ou A* e o *sítio de saída E* (do inglês, *exit*). O sítio P é ocupado assim que a tradução começa pelo $tRNA^{fMet}$, enquanto o sítio A fica posicionado sobre o segundo códon do mRNA e está temporariamente desocupado. O sítio E ligará o tRNA descarregado que está saindo do ribossomo. O alongamento da cadeia polipeptídica é iniciado quando o segundo tRNA aminoacilado entra no sítio A e pareia com o segundo códon por seu braço anticódon. Essa fase requer a presença de dois *fatores de alongamento*, o EF-Tu e o EF-Ts. O EF-Tu se liga ao segundo tRNA que entrará no sítio A e a uma molécula de GTP. Após esse complexo se ligar no sítio A, o GTP é hidrolisado em GDP e o complexo EF-Tu-GDP é liberado. O fator EF-Ts regenera o complexo EF-Tu-GDP em EF-Tu-GTP, deixando-o preparado para participar de uma nova rodada de alongamento.

O contato próximo entre os aminoácidos que estão ligados aos tRNA ocupantes dos sítios P e A favorece a formação de uma ligação peptídica entre o grupamento carboxila do fMet e o grupamento amino do segundo aminoácido, gerando um dipeptídio. Essa reação é catalisada pelo *centro de peptidil transferase* presente na subunidade ribossômica maior. Em seguida, o tRNA no sítio P libera seu aminoácido e o ribossomo desliza 3 nucleotídios ao longo do mRNA no sentido 5′ → 3′, posicionando-se no códon seguinte do mRNA. Esse movimento é chamado de *translocação* e requer a presença do fator G de alongamento (EF-G) e a hidrólise de GTP em GDP. Com a translocação do ribossomo, o tRNA que ocupava o sítio P, passa a ocupar o sítio E e o tRNA que ocupava o sítio A passa a ocupar o sítio P, deixando o sítio A livre para que um novo tRNA aminoacilado possa ocupá-lo. Isso quer dizer que o sítio P é o sítio de entrada apenas para o primeiro tRNA ($tRNA^{fMet}$), enquanto os demais tRNA sempre se ligarão inicialmente no sítio A.

O ciclo de *alongamento* pode, então, ser repetido inúmeras vezes até a formação da cadeia polipeptídica completa. Nesse ciclo, um tRNA aminoacilado reconhece um códon por complementaridade e entra no sítio A. Posteriormente, uma ligação peptídica é formada entre os aminoácidos proximamente localizados nos sítios A e P e o ribossomo desliza-se para o códon seguinte, liberando o sítio A para que um novo tRNA possa se ligar. A cadeia polipeptídica nascente permanece ligada ao ribossomo pelo sítio P. Essas etapas estão resumidas na Figura 2.16.

Para que a tradução seja *finalizada*, um dos 3 *códons de término* (UAA, UAG e UGA) deve ser reconhecido. Entretanto, como não existem tRNA com anticódons que sejam complementares aos códons de término, quando esses códons são encontrados, o sítio A não é preenchido por tRNA, mas, sim, por *fatores de liberação*. Em bactérias, 3 fatores de liberação podem

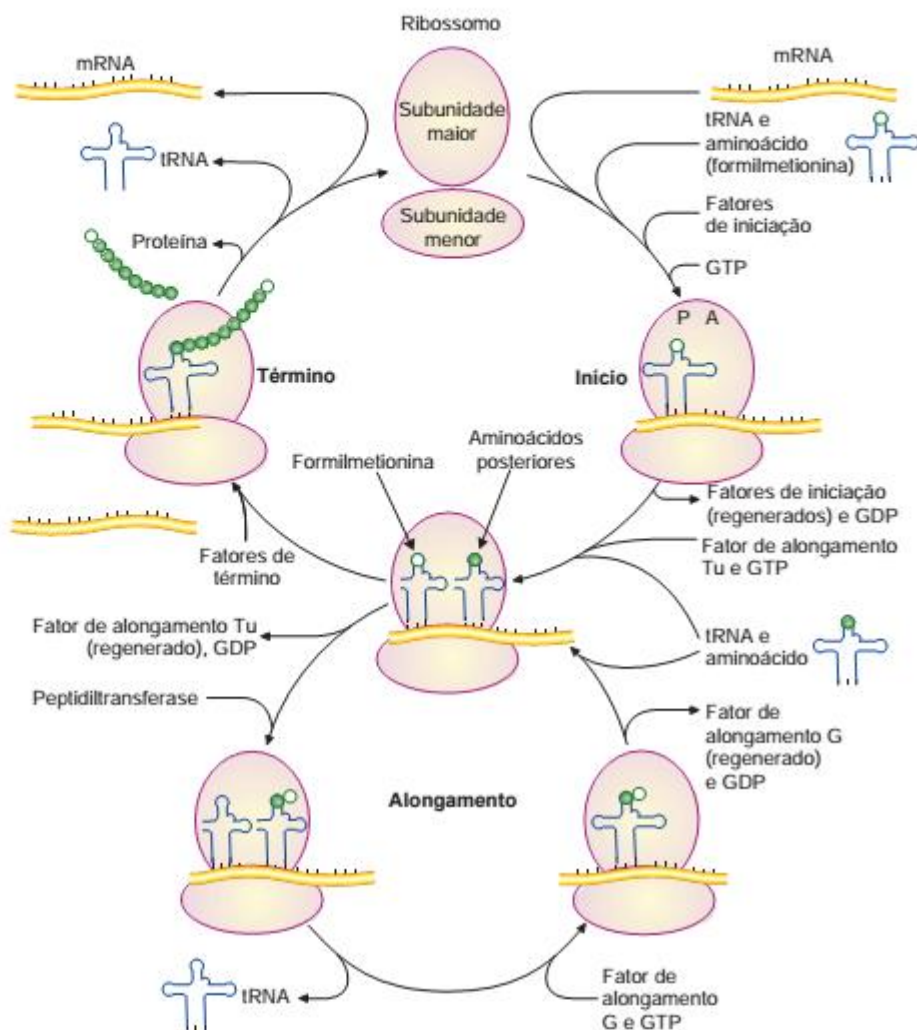



Figura 2.16 Súmula das etapas de início, alongamento e término que compõem a tradução em procariotos.

ser encontrados (RF_1, RF_2 e RF_3). O RF_1 liga-se aos códons UAG e UAA, enquanto o RF_2 liga-se aos códons UAA e UGA. O RF_3 atua como um fator auxiliar, formando um complexo com GTP que se liga ao ribossomo. Mediante a hidrólise de GTP em GDP, os fatores de liberação possibilitam, então, a clivagem do tRNA que está no sítio P e a liberação posterior do polipeptídio. Daí, o aparato traducional é desfeito com a liberação do mRNA do ribossomo e ocorre dissociação do ribossomo em suas duas subunidades que retornam ao *pool* celular para participarem de uma nova tradução.

Tradução em eucariotos

Em caso de genes eucarióticos codificadores de proteínas, após a formação do transcrito de mRNA maduro, a molécula migra para o citoplasma, no qual o processo de síntese proteica acontece. A tradução eucariótica ocorre primordialmente nos ribossomos citoplasmáticos, mas também nas mitocôndrias e nos cloroplastos em plantas. No entanto, embora seja frequentemente aceito que a tradução não aconteça no núcleo eucariótico, evidências recentes sugerem que, em algumas situações, a tradução pode efetivamente ocorrer no núcleo de forma acoplada à transcrição, como é observado em bactérias.

O processo de tradução em eucariotos se assemelha bastante ao de bactérias, mas algumas diferenças importantes podem ser constatadas. Essas peculiaridades possibilitam, igualmente, que estratégias de ação de antibióticos sejam desenvolvidas especificamente para atuarem no

processo de tradução bacteriano, sem que haja riscos de interferência na tradução de genes humanos. Em primeiro lugar, as estruturas ribossômicas nos dois sistemas têm distinções com relação ao tamanho e ao número de moléculas de rRNA presentes, como vimos. Enquanto a subunidade ribossômica maior bacteriana apresenta apenas 2 moléculas de rRNA, em eucariotos são observadas 3 moléculas. Além disso, embora o processo químico de aminoacilação dos tRNA seja o mesmo nos 2 grupos de organismos, o início da tradução apresenta características próprias. Em bactérias, vimos que a tradução é iniciada com o reconhecimento da sequência consenso Shine-Dalgarno pelo rRNA 16S da subunidade ribossômica menor.

Já em eucariotos, o início da tradução acontece pelo reconhecimento da *estrutura cap* na extremidade 5′ da molécula de mRNA pela subunidade ribossômica menor (40S) por meio de proteínas.

Sequência Kozak
Sequência de nucleotídios consenso localizada no mRNA de eucariotos, a qual tem grande importância no início do processo de tradução

A subunidade ribossômica menor se liga, então, ao mRNA nessa região e migra até encontrar o códon de início AUG. Contudo, nem sempre o primeiro códon AUG encontrado no mRNA representa o códon de início da tradução. O reconhecimento do códon AUG correto para início da tradução ocorre mediante a identificação de uma sequência consenso de nucleotídios chamada de sequência Kozak, a qual contém o códon de início (5′-GCCPuCCAUGG-3′, em que Pu é uma purina).

A cauda poli (A) localizada na extremidade 3′ do mRNA, assim como o cap 5′, também exerce papel fundamental no início da tradução.

Essa região é alvo de ligação de proteínas que interagem com proteínas ligadas no cap 5′, formando uma estrutura circular do mRNA. Essa interação é capaz de intensificar a ligação da subunidade ribossômica menor à extremidade 5′ do mRNA.

Existem regiões adjacentes ao segmento codificante que são transcritas, mas não são traduzidas, recebendo o nome de regiões não traduzidas ou UTR (do inglês *untranslated region*). Estas estão presentes tanto anteriormente à região a ser traduzida na região 5′ (5′UTR), como posteriormente a ela, na região 3′ (3′UTR). Apesar de essas regiões não codificarem nenhum aminoácido, elas têm grande importância para a estabilidade do transcrito e para o processo traducional em si. Além da estrutura cap e da cauda poliA, nas regiões 3′UTR e 5′UTR também foram identificados vários elementos aos quais se ligam proteínas capazes de modular a tradução de um mRNA.

Outra diferença importante entre o início da tradução de procariotos e eucariotos é o fato de que em eucariotos o primeiro tRNA a ser acoplado no códon de início transporta uma *metionina não modificada*, de maneira diferente da fMet carreada pelo tRNA iniciador em bactérias. Somado a isso, muitos fatores de iniciação participam da tradução em eucariotos. Esses fatores podem atuar em diferentes instâncias, como na separação das subunidades ribossômicas menor e maior, no reconhecimento da estrutura cap em 5′ do mRNA e na ligação da subunidade menor a essa região, além de participarem como RNA helicases, desespiralizando estruturas secundárias que podem estar presentes no caminho até que o códon de iniciação seja alcançado. Alguns fatores ainda têm a função de direcionar o primeiro tRNA aminoacilado ($tRNA^{Met}$) até o complexo de iniciação.

Com relação às etapas de *alongamento* e *término*, o processo de tradução eucariótico é bem similar ao que acontece em bactérias, com a diferença de que diferentes fatores de alongamento e liberação são utilizados nos dois sistemas. Outro fator que diferencia a tradução em eucariotos da que ocorre em bactérias é a grande *instabilidade* da maioria dos mRNA bacterianos. A degradação do mRNA bacteriano ocorre em poucos minutos depois da tradução. Na realidade, a vida média de um transcrito gênico em *E. coli* é de cerca de 5 min. Curiosamente, a extremidade 5′ de um mRNA bacteriano pode começar a ser degradada antes mesmo de sua extremidade 3′ ter acabado de ser traduzida. No entanto, como as moléculas de mRNA são sintetizadas, traduzidas e degradadas no mesmo sentido (5′ → 3′) e a degradação progride de maneira mais lenta que a tradução, isso possibilita que o polipeptídio seja liberado corretamente. Já os mRNA eucariotos apresentam *estabilidade variável*, podendo ser degradados em algumas horas ou perdurarem por dias, com uma vida média de 5 h. A maior estabilidade dos transcritos gênicos em eucariotos ocorre, em parte, em virtude da associação com proteínas de ligação ao RNA.

É válido ressaltar que tanto em bactérias quanto em eucariotos, vários ribossomos podem se acoplar simultaneamente a uma mesma molécula de mRNA, aumentando a eficiência de tradução.

Polirribossomo
Molécula de mRNA com vários ribossomos a ela ligados

A estrutura formada pela molécula de mRNA que está sendo traduzida e seus vários ribossomos é chamada de polirribossomo. Nesse caso, um ribossomo se liga à extremidade 5′ do mRNA e se desloca em direção à extremidade 3′, liberando espaço físico para que outros ribossomos se liguem consecutivamente à extremidade 5′ da molécula de mRNA.

Tradução de genes mitocondriais

Até o momento, abordamos como o processo de tradução acontece em bactérias e eucariotos. Todavia, você deve estar pensando: e o genoma mitocondrial, como ele é traduzido?

O aparato de tradução de polipeptídios mitocondriais é bem semelhante àquele encontrado nos ribossomos citoplasmáticos. A estrutura dos ribossomos mitocondriais também tem duas subunidades (maior e menor) formadas por rRNA e proteínas. Dentre os 37 genes mitocondriais humanos, 22 codificam moléculas de tRNA mitocondrial e 2 codificam as duas moléculas de rRNA mitocondrial (uma molécula de rRNA 23S, componente da subunidade ribossômica mitocondrial maior, e uma molécula de rRNA de 16S, componente da subunidade ribossômica mitocondrial menor). Os 13 genes mitocondriais restantes codificam polipeptídios que são sintetizados nos ribossomos mitocondriais e têm participação na fosforilação oxidativa. Entretanto, grande parte das subunidades polipeptídicas que atuam no processo de fosforilação oxidativa mitocondrial é codificada por genes nucleares, sendo traduzidas nos ribossomos citoplasmáticos e depois importadas para o interior das mitocôndrias. Além disso, os componentes da maquinaria traducional das mitocôndrias consistem não somente em produtos de genes mitocondriais, como também, em produtos de genes nucleares.

Nesse aspecto, o genoma mitocondrial codifica todas as moléculas de rRNA e tRNA de que precisa durante a tradução, mas depende do genoma nuclear para fornecer os demais componentes, como as 22 aminoacil-tRNA-sintetases e cerca de 80 proteínas ribossômicas. Esses componentes são codificados por genes nucleares, traduzidos por ribossomos citoplasmáticos e, posteriormente, importados para o interior das mitocôndrias, onde exercerão sua função.

Como existem apenas 22 tipos de tRNA mitocondriais humanos, uma mesma molécula de tRNA precisa reconhecer vários códons distintos. Dos 64 códons possíveis, no entanto, quatro não são reconhecidos por tRNA mitocondriais complementares (UAG, UAA, AGA, AGG) e atuam como *códons de término* da tradução mitocondrial.

Vigilância do mRNA

Uma mutação em um mRNA pode favorecer a síntese de uma proteína, cuja função na célula pode ser bruscamente alterada. Em virtude disso, as células eucariotas, assim como as bacterianas, têm mecanismos de *vigilância do mRNA*, os quais possibilitam que determinados mRNA apresentando mutações sejam eliminados antes de gerarem um polipeptídio funcionalmente anormal. Nesse sentido, mRNA que apresentem um códon prematuro de término da tradução anterior ao original em virtude de uma mutação (Capítulo 3) podem ser alvos de um mecanismo chamado de *decaimento de mRNA mediado por mutações nonsense* (do inglês, *nonsense-mediated decay*). Por esse mecanismo, o mRNA mutado é degradado previamente à tradução, levando à perda total de função da proteína em detrimento de uma função anormal da mesma na célula. Em outros casos, o mRNA não apresenta um códon de término, seja em decorrência de uma mutação que levou à conversão do códon de término original em um códon que especifica um aminoácido, seja em virtude de um término precoce da transcrição, deixando o mRNA sem códon finalizador. Nessas situações, o ribossomo se move até o final do mRNA, sem encontrar o códon de término, permanecendo ligado ao mRNA. A ligação do ribossomo ao mRNA evita que esse ribossomo possa ser reutilizado em uma nova tradução, podendo afetar os níveis globais de síntese de proteínas. Em casos como esse, pode ocorrer também degradação do mRNA anormal que não tenha um códon de término.

tmRNA
RNA mensageiro transportador que tem propriedades de mRNA e tRNA, participando da recuperação de ribossomos que permanecem ligados à molécula de mRNA

As bactérias apresentam um mecanismo diferenciado de lidar com essas circunstâncias, nas quais o ribossomo fica estacionado no mRNA sem que a tradução tenha sido finalizada. Quando um ribossomo fica ancorado no mRNA, uma molécula de RNA mensageira transportadora (tmRNA), que apresenta propriedades tanto de mRNA como de tRNA, é levada ao sítio A do ribossomo pelo fator de alongamento EF-Tu. Posteriormente, é criada uma ligação peptídica entre o aminoácido alanina, normalmente transportado pelo tmRNA, e o aminoácido carregado pelo tRNA acoplado no sítio P. Além disso, a parte do tmRNA que funciona como mRNA é traduzida como uma extensão ao polipeptídio inicialmente traduzido com base no mRNA original. Essa tradução acontece até que seja alcançado um códon de término localizado na extremidade 3′ do tRNA. A partir de então, o tmRNA recruta fatores de término apropriados, possibilitando que o ribossomo seja reutilizado em um outro processo de tradução e a proteína com a marca do tmRNA seja liberada. Os aminoácidos adicionais inseridos na proteína original pela tradução do tmRNA possibilitam que a proteína seja subsequentemente reconhecida e degradada por proteases específicas. Os tmRNA ainda não foram identificados em *archae* e no genoma nuclear de eucariotos.

Mutações sinônimas e seu efeito sobre a tradução

Na maioria das vezes, atrelamos a ideia de que a estrutura da proteína depende unicamente da sua sequência de aminoácidos. Por esse motivo, é esperado que mutações sinônimas caracterizadas por não ocasionarem a troca de um aminoácido não teriam, na maioria dos casos, uma consequência fenotípica, a não ser que estejam presentes em elementos participantes do processo de recomposição dos éxons. No entanto, evidências crescentes apontam que mesmo que o códon substituto especifique o mesmo aminoácido que o original, essa modificação pode, sim, ter efeitos nocivos (Capítulo 3). Embora o código genético seja degenerado, códons que codificam um mesmo aminoácido são utilizados de maneira desigual pelas células. Desse modo, códons sinônimos podem apresentar *velocidades de tradução distintas* em decorrência dos diferentes níveis de abundância dos tRNA. Isso poderia influenciar o tempo disponível para a dobradura da proteína que ocorre após a tradução e, assim, alterar a estutura e a função proteicas.

Modificações pós-traducionais

As proteínas podem ser organizadas em 4 níveis distintos (*primário*, *secundário*, *terciário* e *quaternário*). O nível *primário* é representado pela sequência linear dos aminoácidos que a compõem, podendo variar extensamente em comprimento. As interações entre os aminoácidos possibilitam que a estrutura primária se dobre em uma *estrutura secundária*. Esse segundo nível de organização molecular das proteínas representa a trajetória espacial do arcabouço polipeptídico. As principais estruturas secundárias encontradas são representadas pela α*-hélice* e a *folha* β*-pregueada*, que são estabilizadas pela formação de pontes de hidrogênio entre grupamentos de aminoácidos diferentes. A α-hélice tem formato de cilindro rígido e é estabilizada pela formação de pontes de hidrogênio entre o grupamento CO de um aminoácido com o grupamento NH de outro aminoácido localizado a 4 unidades adiante, sendo comumente encontrada em fatores que se ligam ao DNA. A folha β pregueada tem formato achatado rígido e ocorre quando um polipeptídio se dobra sobre si mesmo repetidamente e os segmentos paralelos são estabilizados pela formação de pontes de hidrogênio entre ligações peptídicas vizinhas.

Os componentes estruturais secundários de um polipeptídio têm a capacidade de se dobrar ainda mais até formar uma estrutura tridimensional chamada de *estrutura terciária*. É a estrutura terciária que confere a atividade biológica do polipeptídio. Nessa conformação, em geral, os aminoácidos com cadeias laterais hidrofóbicas ficam localizados no interior da proteína, ao passo que os aminoácidos hidrofílicos ficam voltados para a superfície das proteínas em contato com o citoplasma. A estabilização dessa conformação pode ser mantida por ligações covalentes relativamente fracas e pontes dissulfídricas entre cisteínas. Técnicas como a difração de raios X, a cristalografia de proteínas e a ressonância magnética nuclear podem determinar a estrutura tridimensional das proteínas. Além disso, predizer de modo acurado a estrutura de uma proteína também tem sido referida como o “Santo Graal” da Bioinformática Estrutural.

Devemos lembrar que muitas proteínas são formadas por mais de uma cadeia polipeptídica, de modo que dizermos que um gene codifica uma proteína pode não ser verdade para todas as situações. A interação entre duas ou mais cadeias polipeptídicas pode levar à formação de um arranjo complexo de uma proteína multimérica. Essa estrutura em nível *quaternário* pode ser encontrada em proteínas que têm várias subunidades, como é o caso das RNA polimerases e da hemoglobina.

Após a tradução de uma proteína bacteriana ou eucariota, podem ocorrer modificações na estrutura da proteína precursora. O polipeptídio recém-traduzido ainda está inativo e, para exercer sua função, a proteína deve inicialmente se *dobrar* até adquirir sua estrutura terciária correta. Algumas proteínas se dobram espontaneamente, enquanto outras necessitam do auxílio de proteínas moldadoras, chamadas de chaperonas. Essa primeira modificação pós-traducional deve acontecer em todas as proteínas sintetizadas para que elas possam exercer sua função. No entanto, outras modificações subsequentes também podem ser imprescindíveis para que a proteína adquira sua conformação tridimensional correta ou se torne funcionalmente ativa, incluindo a remoção do primeiro aminoácido especificado, a modificação química de aminoácidos específicos ao longo da cadeia polipeptídica e a clivagem de um segmento do produto proteico primário.

Chaperonas
Proteínas que auxiliam na dobradura de outras proteínas

Com relação à remoção do primeiro aminoácido especificado, é sabido que algumas proteínas podem sofrer excisão do grupamento formil ou do aminoácido metionina inicial localizado na extremidade aminoterminal da proteína por enzimas específicas. A metionina, ou formilmetionina, especificada pelo códon de início AUG, é utilizada apenas para sinalizar o início da tradução; depois que a tradução foi finalizada, sua presença não é mais tão essencial para alguns polipeptídios. Contudo, modificações químicas também podem ocorrer em outros aminoácidos específicos após a tradução, incluindo hidroxilação, fosforilação, metilação, acetilação, glicosilação ou a adição de carboidratos e lipídios. A inserção de grupamentos químicos nesses aminoácidos muitas vezes é crucial para a atividade biológica da proteína.

Além da remoção do primeiro aminoácido e das modificações químicas, clivagens proteolíticas do produto primário proteico também podem ser necessárias após a tradução, produzindo um polipeptídio maduro menor. Nesse contexto, algumas proteínas têm em sua estrutura uma sequência de aminoácidos chamada de *sequência de sinalização*. Esta tem a finalidade de funcionar como um sinal para seu transporte até compartimentos celulares específicos (retículo endoplasmático, mitocôndrias, núcleo etc.), possibilitando que a proteína seja direcionada para seu local correto de atuação. Assim que o destino é alcançado, a sequência de sinalização pode ser, então, removida pela ação de proteases. Esse evento é comum em proteínas plasmáticas, hormônios, neuropeptídios e fatores de crescimento. Um exemplo de polipeptídio primário que sofre clivagem proteolítica é representado pela síntese de *insulina*, a qual é inicialmente traduzida como um produto primário maior inativo, chamado de *pré-proinsulina*, que apresenta uma sequência de sinalização em sua estrutura de aminoácidos. Essa sequência é essencial para que o polipeptídio atravesse a membrana celular. Tão logo o objetivo seja alcançado, a sequência de sinalização é descartada e o polipeptídio é transformado em *proinsulina*. Esta apresenta uma sequência central (*peptídio conector*) que parece ter papel importante na manutenção da conformação das cadeias A e B da insulina madura. Esse peptídio conector também é clivado por ação proteolítica e as *cadeias A e B* formam pontes dissulfídricas entre si, gerando a forma madura e biologicamente ativa da insulina.

É válido mencionar que alguns produtos primários da tradução em eucariotos e procariotos apresentam sequências curtas de aminoácidos que têm a capacidade de se autorremoverem do polipeptídio recém-sintetizado. Essas sequências recebem o nome de *inteínas*, em analogia aos íntrons que são removidos no processamento do RNA. Após a remoção das inteínas, as "*exteínas*" a elas adjacentes são unidas para que a proteína se torne ativa. Uma das primeiras inteínas conhecidas encontra-se na proteína RecA, envolvida na recombinação e no reparo do DNA na bactéria *Mycobacterium tuberculosis* que causa a tuberculose.

■ Regulação da expressão gênica

Em todo organismo vivo, existem genes que necessitam ser expressos de modo basal a todo o momento, por participarem da sustentação do funcionamento da célula e do organismo como

Genes housekeeping
Genes que regem funções celulares básicas para o organismo e são expressos em níveis basais constantes

um todo. Esses genes recebem a denominação *genes de manutenção* ou "genes housekeeping", sendo representados por genes que codificam enzimas como a DNA e a RNA polimerase, as quais são usadas rotineiramente pelas células para a manutenção basal. Todavia, outros genes mais especializados só precisam ser expressos, sob a forma de uma molécula de RNA ou proteína, em determinada circunstância temporal ou espacial. Assim, para controlar quais genes devem se expressar em dado momento, todos os organismos vivos têm a habilidade de regular a expressão de seus genes.

Esse controle possibilita que aqueles genes cujos produtos não são necessários naquela ocasião sejam desligados, e outros, que precisam estar ativos, sejam estimulados a se expressarem.

Quando a regulação da expressão gênica favorece a expressão de um gene, dizemos que a regulação é *positiva*. De modo contrário, quando o processo de regulação inibe a expressão de um gene dizemos que o processo é *negativo*.

No genoma de um organismo vivo, podemos encontrar dois grupos de genes: os *genes estruturais*, que codificam proteínas com função estrutural ou enzimática, e os *genes reguladores*, cujos produtos gênicos, em forma de RNA ou proteínas, interagem com outras sequências de DNA reguladoras (elementos reguladores), modulando as taxas de transcrição e tradução dessas sequências. Os elementos reguladores consistem em sequências de DNA que geralmente não são transcritas e funcionam como sítio de ligação de proteínas produzidas pelos genes reguladores, modulando a expressão de genes-alvo.

A regulação da expressão gênica em procariotos e eucariotos pode acontecer em diferentes níveis no caminho molecular que vai desde a molécula do DNA até o produto proteico final.

O ponto de regulação primário em eucariotos ocorre ainda em *nível de DNA* pela *alteração dinâmica da estrutura da cromatina*. Nesse nível, modificações químicas na molécula de DNA ou nas caudas das histonas determinam que sequências gênicas estarão disponíveis para que a maquinaria transcricional possa se ligar. O segundo nível de regulação comum a procariotos e eucariotos acontece durante a *transcrição* e visa à restrição da síntese de produtos gênicos precocemente objetivando primordialmente a economia no gasto de energia. Um terceiro nível pelo qual um gene pode ser regulado acontece no *processamento do mRNA* eucariótico. Como visto anteriormente, o amadurecimento do pré-mRNA envolve modificações (capeamento na extremidade 5′, poliadenilação na extremidade 3′ e remoção dos íntrons) que influenciam a estabilidade da molécula de mRNA madura e a capacidade de tradução dessa molécula. Com isso, mecanismos reguladores podem atuar durante o processamento do mRNA, modulando a expressão gênica. Na realidade, a estabilidade do RNA determina diretamente a quantidade de proteína que será produzida. Quanto maior for a escala de degradação de um mRNA, menor será a taxa de síntese de polipeptídios. Na *tradução*, vários fatores também podem influenciar a taxa de síntese de uma proteína em procariotos e eucariotos e, consecutivamente, influenciar a expressão do gene que a codifica. Por exemplo, se vários ribossomos se ligam a um mesmo mRNA (poliribossomo), a taxa de síntese aumentará consideravelmente. O último momento no qual um gene pode ser regulado é em nível *pós-traducional*. Vimos que as proteínas sofrem modificações após sua tradução, as quais possibilitam que elas se tornem funcionalmente ativas. Assim, fatores que alterem essas modificações pós-traducionais podem regular um gene em um passo mais tardio da expressão gênica.

Do mesmo modo como descrevemos os mecanismos de transcrição e tradução de maneira independente em bactérias e eucariotos, também o faremos para a regulação da expressão gênica.

Regulação da expressão gênica em procariotos

O genoma bacteriano carrega consigo as informações biológicas para sintetizar uma gama de produtos gênicos diferenciados. Contudo, as bactérias utilizam o *princípio da economia máxima de energia* e da *flexibilidade bioquímica* para regular seus genes, de modo que somente serão expressos aqueles produtos gênicos necessários para cada condição presente no ambiente. Exemplificando este conceito, a *E. coli* apresenta vários genes que codificam

enzimas metabolizadoras de diferentes tipos de açúcar. Esses genes poderiam ser expressos a todo o momento, mas isso acarretaria um *gasto desnecessário de energia*, visto que nem todos os tipos de açúcar estão disponíveis concomitantemente no ambiente. Sendo assim, a *E. coli* expressa apenas os genes específicos para os açúcares presentes naquele momento, desligando os demais genes dispensáveis. Caso haja mudança na condição ambiental, a *E. coli* apresenta a flexibilidade bioquímica de redefinir os seus padrões de expressão gênica, desligando genes que estavam ativos e ligando genes que estavam inativos.

Conforme descrevemos, muitos genes funcionalmente relacionados encontram-se aglomerados no genoma bacteriano de modo contíguo, estando sob o controle de um único promotor e sendo transcritos juntos em uma mesma molécula de RNA. Ao conjunto de genes estruturais transcritos juntos, seu promotor e suas sequências reguladoras adicionais é dado o nome de **óperon**. Na Figura 2.17, podemos evidenciar uma estrutura clássica de um óperon, formado por um grupo de 3 genes estruturais ordenados lado a lado (A, B e C), seu promotor comum localizado antes do primeiro gene estrutural e as sequências reguladoras. Entre as sequências reguladoras, evidencia-se uma chamada de *operador* que é contígua aos genes estruturais e normalmente se sobrepõe ao promotor e, às vezes, também ao primeiro gene estrutural. Alguns óperons apresentam vários operadores. Fora do óperon, podemos encontrar ainda um gene regulador que apresenta seu próprio promotor e codifica pequena proteína reguladora. A proteína codificada pelo gene regulador pode se ligar ao operador, afetando o funcionamento transcricional do óperon.

Óperon
Conjunto de genes estruturais bacterianos, seu promotor comum e outras sequências reguladoras flanqueadoras que controlam a expressão dos genes estruturais

Óperon indutível
Óperon no qual a transcrição dos genes estruturais está normalmente desligada e uma mudança deve acontecer para que a transcrição seja ativada

Óperon repressível
Óperon no qual a transcrição dos genes estruturais está normalmente ligada e uma mudança deve acontecer para que a transcrição seja desativada

Os óperons podem ser de dois tipos: **óperons indutíveis** e **óperons repressíveis**. Nos indutíveis, a transcrição está inativa e uma mudança deve acontecer para estimular a transcrição. Nos óperons repressíveis, a transcrição está ativa e uma mudança deve acontecer para inibir a transcrição. Além desses diferentes tipos de óperons, também podemos evidenciar dois tipos de controle da expressão gênica (*negativo* e *positivo*). No *controle negativo*, o produto do gene regulador é uma proteína reguladora *repressora* que se liga ao operador e *inibe* a transcrição dos genes estruturais. No controle positivo, o produto do gene regulador é uma proteína reguladora *ativadora*, que é necessária para *induzir* a transcrição dos genes estruturais.

Controle negativo de regulação bacteriana da expressão gênica

Em um *óperon indutível negativo*, a expressão do *gene regulador* gera um *repressor* que se liga normalmente ao *operador*. Como o operador está sobreposto à *região promotora*, a ligação do repressor impede fisicamente a ligação da RNA polimerase ao promotor, impossibilitando a transcrição dos genes estruturais presentes no óperon. Para que a transcrição seja ativada, é necessário que uma pequena molécula *indutora* se ligue ao *repressor*, causando *mudança conformacional alostérica* do repressor. Com isso, o repressor não consegue mais se ligar ao operador e a transcrição pode transcorrer normalmente, já que a RNA polimerase tem passe

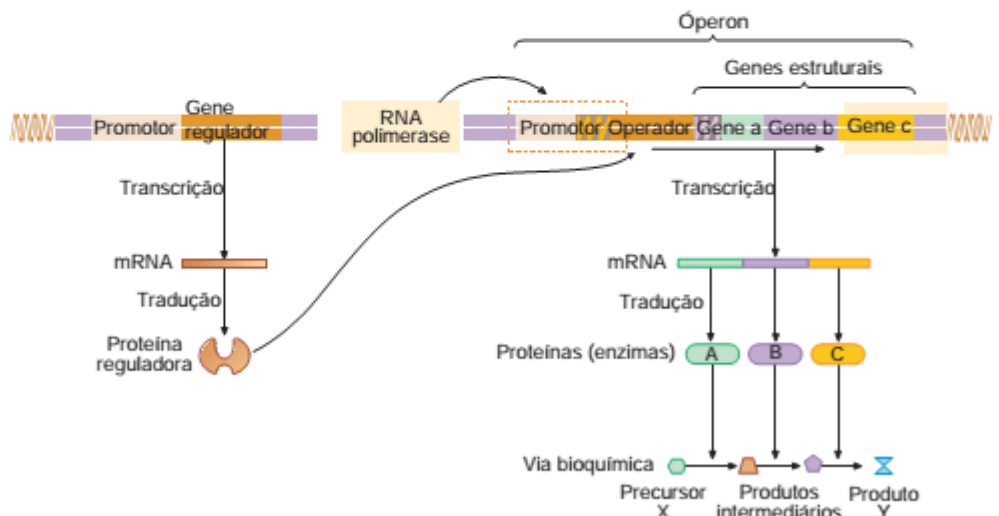



Figura 2.17 Estrutura clássica de um óperon bacteriano, formado por três genes estruturais (A, B e C), o promotor comum e as sequências reguladoras.

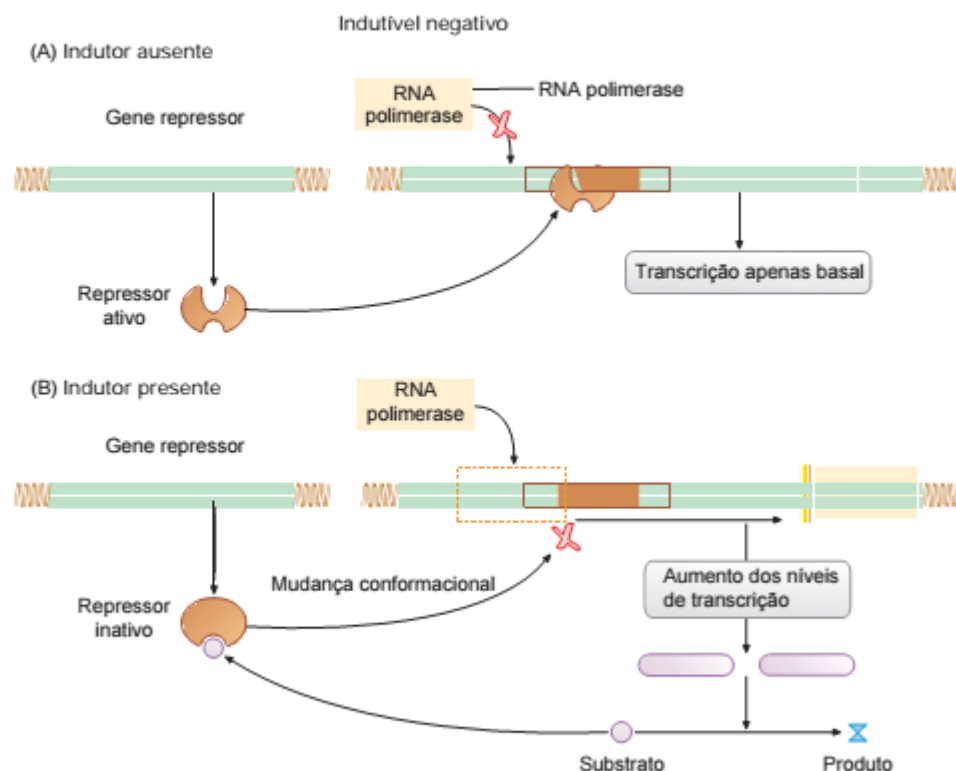



Figura 2.18 Modelo de óperon indutível negativo. Na ausência (A) e presença (B) do elemento indutor

livre para se ligar ao promotor. Podemos dizer então que o óperon indutível negativo está normalmente desligado e após uma indução, a transcrição se torna ativa (Figura 2.18).

Os *óperons indutíveis negativos*, em geral, controlam proteínas que têm função de quebrar moléculas, de modo que sua presença não é necessária na ausência de um substrato apropriado. Um exemplo clássico de óperon indutível em bactérias é representado pelo *óperon da lactose* (óperon Lac) em *E. coli*. A lactose é um dos principais carboidratos presentes no leite e pode ser metabolizada por bactérias *E. coli* presentes no tubo digestivo de mamíferos. Para utilizar a lactose como fonte de carbono e energia, a *E. coli* a transporta para o interior da célula e depois a hidrolisa em glicose e galactose. Esse processo conta com a participação de 3 enzimas codificadas por genes estruturais dispostos consecutivamente: a *permease* (codificada pelo gene Lac Y), a β-*galactosidase* (codificada pelo gene Lac Z) e a *transacetilase* (codificada pelo gene Lac A). A permease é responsável pelo transporte da lactose para o interior da célula, a β-galactosidase atua na quebra da lactose em glicose e galactose e a transacetilase não têm função bem estabelecida. Os genes LacY, LacZ e LacA estão dispostos de *modo ordenado*, lado a lado, formando o óperon Lac, apresentando o mesmo promotor (LacP) e operador (LacO). Anteriormente ao óperon Lac há um gene repressor chamado de LaI que regula a expressão dos 3 genes estruturais. A ligação do repressor codificado pelo gene LacI ao operador bloqueia o acesso da RNA polimerase ao promotor e inibe a transcrição dos genes estruturais LacZ, LacY e LacA. É nesse estado de inativação transcricional que as células se encontram quando não há lactose no meio, no qual a *E. coli* cresce. No entanto, a repressão nunca desliga completamente a transcrição do óperon Lac, possibilitando a síntese de permease, β-galactosidase e transacetilase em níveis baixos.

Na presença de lactose, as pequenas quantidades de permease, β-galactosidase e transacetilase presentes são suficientes para converter parte da lactose em *alolactose*. A alolactose atua como um indutor e se liga ao repressor LaI, mudando sua conformação e impedindo que ele se ligue ao operador. A partir daí, o promotor fica livre para a interação com a RNA polimerase. Com isso, os 3 genes estruturais podem ser transcritos e posteriormente traduzidos, metabolizando apropriadamente a lactose presente no meio. Caso a lactose acabe, novamente o óperon Lac será desligado, pois o repressor se dissociará do indutor e se ligará novamente ao promotor. Um dado interessante que será comentado mais adiante é o fato de os genes

estruturais no óperon Lac serem transcritos em altos níveis somente na presença de lactose e ausência de glicose. Na presença de glicose, ela é utilizada como fonte preferencial de energia pela célula em vez da lactose.

Em um *óperon repressível negativo*, a transcrição transcorre normalmente, mas em decorrência de uma mudança no ambiente, os genes estruturais presentes no óperon devem deixar de transcrever. A proteína reguladora no óperon repressível é um repressor produzido sob uma forma *inativa* que é incapaz de se ligar ao operador. Como o operador está desocupado, a RNA polimerase pode se ligar à região promotora e efetuar a transcrição dos genes estruturais adjacentes. Para inativar a transcrição, uma pequena molécula *correpressora* se liga ao repressor, causando mudança conformacional que possibilite que o repressor, agora *ativo*, possa se ligar ao operador. Com a ligação do repressor ao operador, a RNA polimerase não pode mais interagir com o promotor e a transcrição dos genes estruturais é impedida. Da mesma maneira que os óperons indutíveis, os repressíveis também visam à minimização do gasto de energia, produzindo apenas as proteínas necessárias em dado momento (Figura 2.19).

Os óperons repressíveis normalmente controlam proteínas que participam da biossíntese de moléculas. Como essas proteínas são rotineiramente necessárias para o funcionamento celular, o óperon repressível está naturalmente ligado e só é desligado quando uma quantidade satisfatória dos produtos dos genes estruturais é alcançada. A partir daí, qualquer excesso no funcionamento do óperon contribuiria somente para um desperdício de energia.

Um exemplo de óperon repressível é representado pelo *óperon do triptofano* (óperon trp) em *E. coli*. Esse óperon é formado por 5 genes estruturais (trpE, trpD, trpC, trpB e trpA), que produzem 3 enzimas, 2 delas constituídas por 2 cadeias polipeptídicas, que participam da conversão do corismato no aminoácido triptofano. Em uma região anterior aos genes estruturais, está o promotor trp. Mais distante do óperon trp, está o gene regulador trpR, que codifica um repressor trp. Este não consegue inicialmente se ligar ao operador quando os níveis de triptofano ainda estão baixos e o óperon trp permanece ligado, ou seja, os genes estruturais estão sendo transcritos e o corismato pode ser convertido em triptofano. Quando a quantidade de triptofano aumenta no meio, a conversão de corismato a triptofano não se torna mais necessária. Com isso, o triptofano atua como ativador, ligando-se ao repressor e causando mudança alostérica.

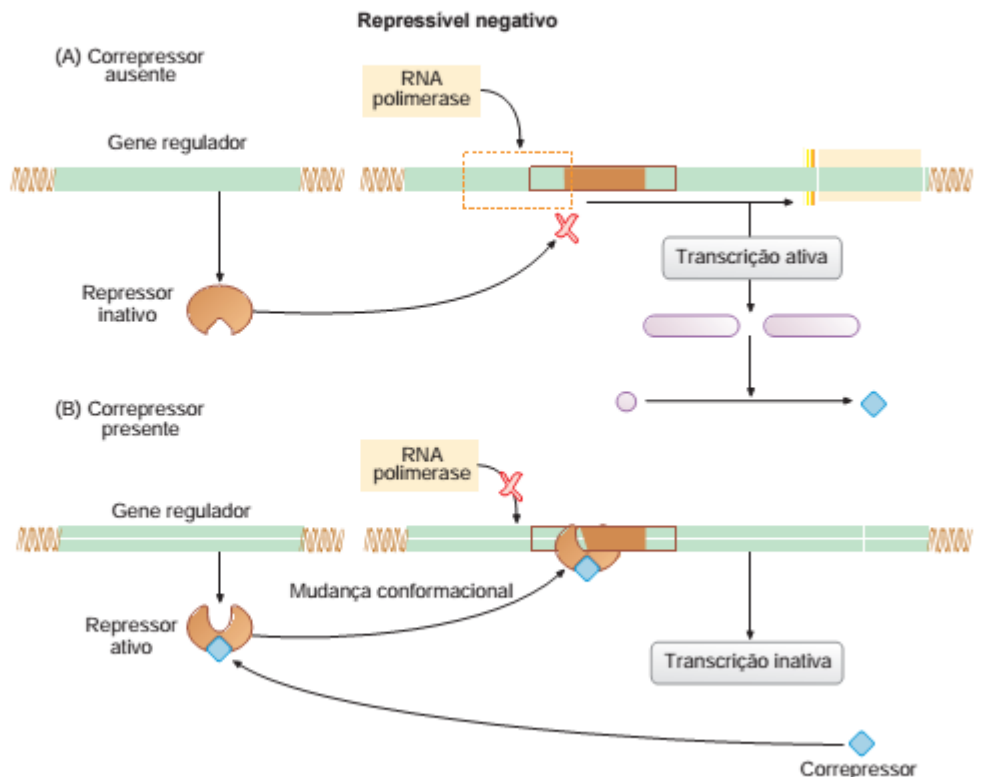



Figura 2.19 Modelo de óperon repressível negativo. Na ausência (A) e presença (B) da molécula correpressora

O repressor pode, então, se ligar ao operador, impedindo a ligação da RNA polimerase ao promotor, desligando, assim, a transcrição dos genes estruturais.

Controle positivo de regulação bacteriana da expressão gênica

Os modelos de óperons indutíveis e repressíveis abordados até o momento têm como ponto comum o *controle negativo* da expressão gênica, representado pela atuação de uma proteína reguladora *repressora*. Vamos considerar agora modelos de óperon, nos quais o controle é *positivo*, envolvendo a participação de uma proteína reguladora *ativadora* capaz de estimular a transcrição. Essa proteína geralmente se liga ao DNA em um sítio diferente do operador, induzindo a transcrição.

Em um *óperon indutível positivo*, os genes estruturais estão normalmente desligados. Nessa situação, a proteína reguladora produzida pelo gene regulador, agora chamada de *ativador*, encontra-se em um estado *inativo*. Para que a transcrição seja iniciada, é necessário que um indutor se ligue à proteína reguladora ativadora, tornando-a *ativa*. O indutor é representado pelo precursor da reação controlada pelo óperon. Sendo assim, visando ao mesmo princípio de economia de energia que os demais óperons descritos, as enzimas codificadas pelos genes estruturais somente são codificadas quando o substrato está presente no meio (Figura 2.20).

De modo contrário ao modelo de óperon indutível positivo, a transcrição de genes estruturais em um *óperon repressível positivo* normalmente está ativa e precisa ser reprimida por uma modificação da condição do meio. Nessa situação, a proteína reguladora ativadora é produzida em uma conformação ativa que está apta a se ligar ao DNA e incitar a transcrição dos genes estruturais. Quando há necessidade de inibir a transcrição por uma alteração do meio, uma substância repressora se liga ao ativador, transformando-o em uma forma inativa. Nessa nova forma, o ativador inativo não consegue mais se ligar à sequência de DNA-alvo e a transcrição dos genes estruturais não pode mais ser estimulada. A substância repressora é representada pelo produto da reação controlada pelo óperon, de modo que quando o produto está disponível, não há necessidade de transcrever genes responsáveis pela sua síntese (Figura 2.21).

Um modelo de *controle positivo* pode ser representado pela situação na qual a glicose impede a indução do óperon Lac e de outros óperons que controlam a síntese de enzimas envolvidas no

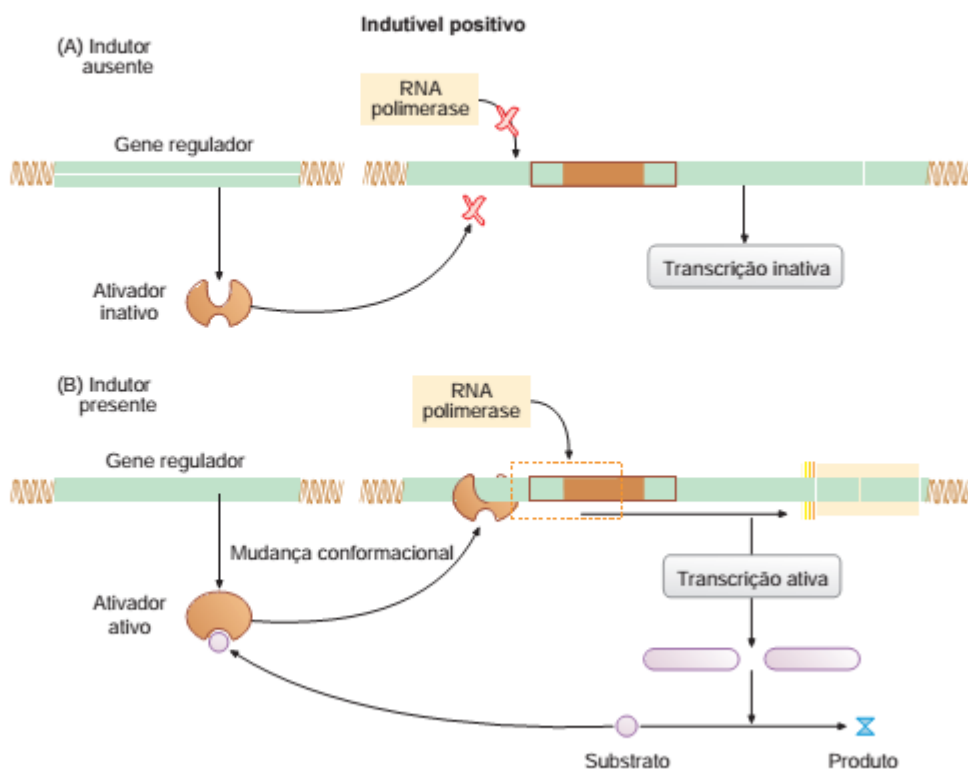



Figura 2.20 Modelo de óperon indutível positivo. Na ausência (A) e presença (B) do indutor

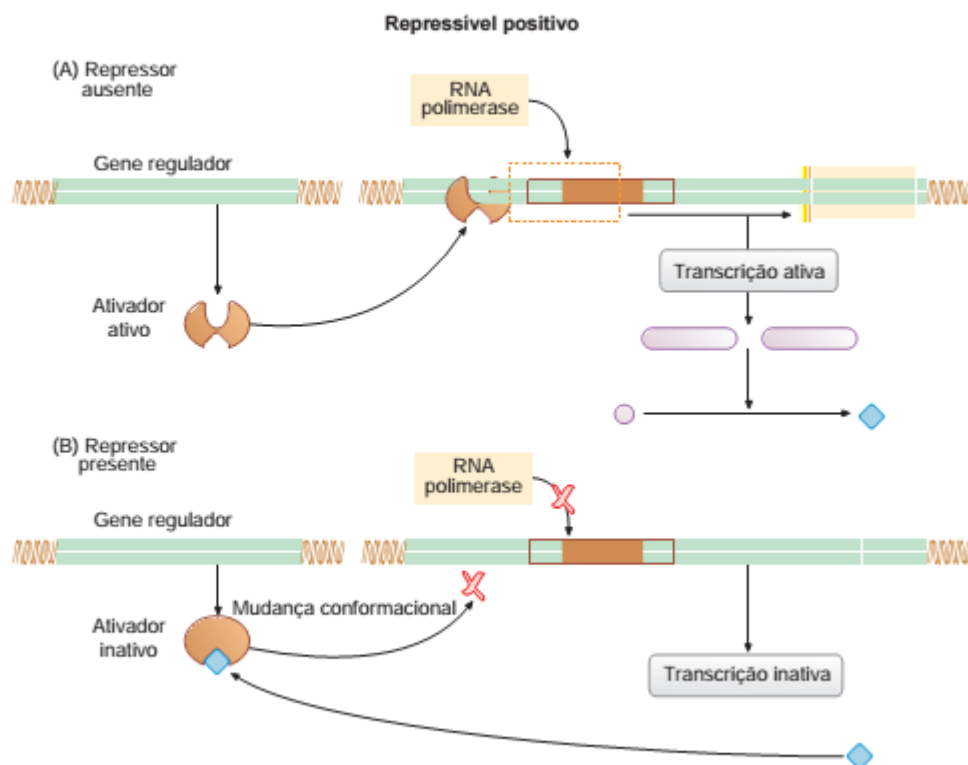



Figura 2.21 Modelo de óperon repressível positivo. Na ausência (A) e presença (B) do repressor.

catabolismo de carboidratos. Quando a glicose está presente no meio, ela é preferencialmente utilizada como fonte de energia em detrimento de outros açúcares. Isso acontece porque durante o metabolismo da glicose ela sofre glicólise sem ser modificada posteriormente, o que representa uma economia de gasto energético pela célula em relação à utilização de outros açúcares. À repressão do metabolismo de outros açúcares na presença de glicose é dado o nome de repressão catabólica. Esse mecanismo de controle positivo envolve a atuação de uma proteína ativadora do catabolismo (CAP). Essa proteína dimérica tem a capacidade de se ligar a uma região que antecede a sequência promotora do óperon Lac. A RNA polimerase somente é capaz de se ligar eficientemente à região promotora se a proteína CAP estiver ligada ao DNA. Caso o sítio de ligação da CAP esteja vago, o óperon será transcrito em taxas pequenas. No entanto, a CAP somente se liga a esta região na presença de AMP cíclico (cAMP), que deriva do ATP por meio de uma reação química catalisada pela adenilato ciclase. Quanto maior for a concentração de adenilato ciclase, maior será a concentração de cAMP. No entanto, a glicose inibe a adenilato ciclase, de modo que quanto maior a quantidade de glicose, menor será a quantidade de cAMP. Com isso, se há muita glicose na célula, a quantidade disponível de cAMP será pequena e níveis reduzidos do complexo cAMP-CAP poderão se ligar ao DNA. Consequentemente, a RNA polimerase terá pouca afinidade pelo promotor Lac e os genes estruturais do óperon Lac serão transcritos em níveis baixos. Assim, se a glicose e a lactose estiverem presentes concomitantemente no meio, a célula utilizará a glicose primeiramente. Por outro lado, se os níveis de glicose são baixos, grandes quantidades de cAMP estarão disponíveis, aumentando também a taxa de ligação do complexo cAMP-CAP ao DNA. Por conseguinte, a eficiência de ligação da RNA polimerase ao promotor será aumentada, assim como a capacidade de transcrição dos genes estruturais Lac, que pode sofrer um acréscimo em 50 vezes.

Repressão catabólica
Controle gênico em alguns óperons bacterianos, nos quais a glicose é utilizada preferencialmente em detrimento de outros açúcares, os quais têm seu metabolismo reprimido pela presença de glicose

Mecanismo de regulação gênica bacteriana por atenuação

Os modelos de regulação discutidos até o momento controlam a expressão gênica no *início da transcrição*. No entanto, alguns óperons têm a habilidade de regular a transcrição em seu *prosseguimento*, processo que é chamado de atenuação. Nesse mecanismo adicional de controle da expressão gênica que acontece em óperons que codificam enzimas envolvidas na biossíntese

Atenuação
Tipo de controle da expressão gênica em alguns óperons bacterianos, nos quais a transcrição é finalizada prematuramente antes da RNA polimerase atingir os genes estruturais

de aminoácidos, a transcrição começa no sítio de início, mas é finalizada prematuramente antes de a RNA polimerase alcançar os genes estruturais.

Um exemplo de óperon que é regulado por atenuação é o *óperon trp* em *E. coli*. Esse óperon apresenta uma situação atípica em relação à maioria dos óperons, pois é regulado tanto por repressão/desrepressão no sítio operador quanto por atenuação da transcrição. Quando o triptofano está aumentado no meio em que a bactéria está sendo cultivada, o término prematuro da transcrição no óperon trp por atenuação acontece em um sítio localizado na extremidade 5′UTR. Mas como isso acontece?

A região 5′UTR presente no transcrito nascente do óperon trp apresenta uma sequência de 162 nucleotídios. Essa sequência líder tem 4 regiões (1, 2, 3 e 4), sendo que a região 1 é complementar à 2, a região 2 é complementar à 3 e a região 3 é complementar à 4. A complementaridade entre essas regiões possibilita que elas interajam entre si na molécula de RNA nascente por meio de pontes de hidrogênio, formando estruturas secundárias em forma de grampo. É válido notar que embora a região 2 possa parear com as regiões 1 e 3, o pareamento somente pode acontecer com uma dessas regiões de cada vez. Na realidade, como podemos observar na Figura 2.22, existem duas possibilidades de formação de estruturas secundárias para a sequência líder do óperon trp. Na primeira possibilidade, a região 1 pareia com a região 2 e a região 3 pareia com a região 4. Na segunda, ocorre pareamento entre as regiões 2 e 3, deixando as regiões 1 e 4 não pareadas. Um dado importante é que após o grampo formado pelo pareamento entre as regiões 3 e 4, há uma sequência rica em uracilas que faz com que essa estrutura se assemelhe muito aos sinais de término da transcrição bacterianos.

Quando o nível de triptofano no meio é baixo, ocorre a formação de grampo entre as regiões 2 e 3. Como o grampo entre essas regiões não é seguido de uma sequência de uracilas, essa estrutura não é reconhecida como um sinal de término e a RNA polimerase continua a transcrever os genes estruturais localizados mais adiante. As enzimas que sintetizam o triptofano podem então ser produzidas, aumentando os níveis de triptofano disponíveis. Quando os níveis de triptofano se tornam elevados, o pareamento ocorre simultaneamente entre as regiões 1 e 2 e as regiões 3 e 4. O pareamento entre as regiões 3 e 4 é lido pela RNA polimerase como um sinal de término da transcrição. Com isso, o triptofano deixa de ser sintetizado, já que os genes estruturais não serão transcritos. Pelo fato de a estrutura formada pelo pareamento 3 + 4 ser capaz de controlar a expressão dos genes estruturais, ela é chamada de *atenuador*. Além do óperon trp, a regulação de um óperon por atenuação também pode ser encontrada em outros cinco óperons (ter, ilv, leu, fen e his).

Entretanto, como os níveis de triptofano podem gerenciar a formação de um ou outro tipo de estrutura secundária (grampos 1/2, 3/4 ou grampo 2/3)?

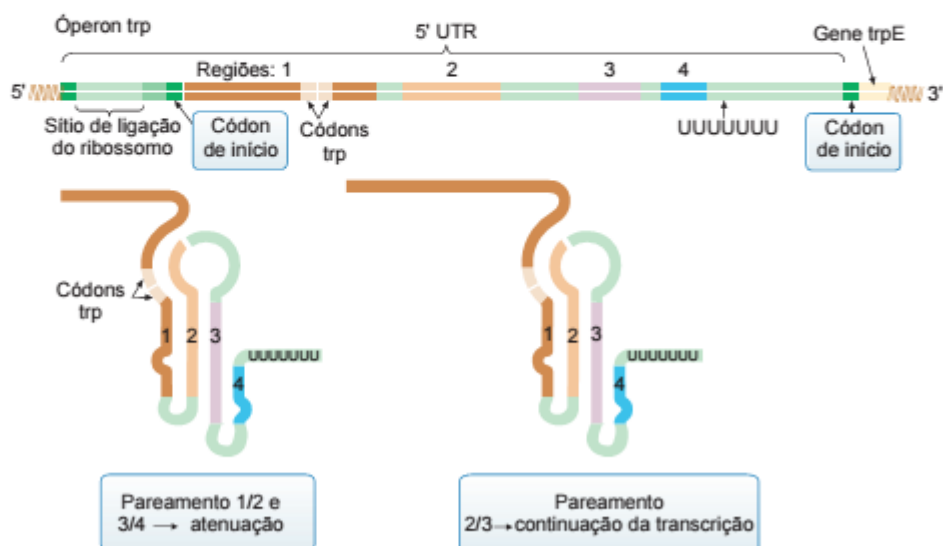



Figura 2.22 Modelo de regulação da expressão gênica bacteriana por atenuação.

Como a transcrição e a tradução em procariotos ocorre em um mesmo espaço físico e de maneira acoplada, enquanto a região 2 está sendo transcrita, a região 1 já está sendo traduzida. Como o ribossomo está ocupando o espaço da região 1 para traduzi-la, ela não pode parear com a região 2. Paralelamente, a região 3 começa a ser transcrita e o ribossomo encontra os códons que especificam triptofano na região 1. Se os níveis intracelulares de triptofano estão baixos, os tRNA que os transportam estão indisponíveis e o ribossomo irá parar a tradução neste ponto, aguardando a chegada de um tRNA carregando um triptofano ($tRNA^{trp}$). Simultaneamente, a região 3 vai sendo transcrita pela RNA polimerase, já que a parada do ribossomo na região 1 não afeta a eficiência da transcrição. Como o ribossomo fica estacionado na região 1, a região 2 após ser transcrita fica livre para formar um grampo com a região 3. O grampo 2/3 impede a formação do grampo 3/4 (atenuador) e como o primeiro não é lido como um sinal de término da transcrição, a transcrição continua normalmente. O triptofano será, então, reposto a partir da expressão dos genes estruturais.

Quando os níveis de triptofano intracelulares estão altos, o ribossomo não ficará estacionado na região 1, pois existirão $tRNA^{trp}$ disponíveis para parearem com os códons de triptofano. Dessa forma, a tradução pode acontecer simultaneamente à transcrição. Quando a região 3 tiver sido transcrita pela RNA polimerase, o ribossomo que ainda está traduzindo a região 2, impede fisicamente que as regiões 2 e 3 pareiem entre si formando um grampo 2/3. Ao mesmo tempo, a RNA polimerase se desloca ao longo do DNA e transcreve a região 4. Como a região 3 não pode parear com a região 2, por empecilho do ribossomo, ela pareia com a região 4, formando o grampo 3/4. Esse grampo seguido das uracilas é reconhecido como um sinal de término pelo ribossomo e a tradução termina logo depois da região 4. Dessa maneira, os genes estruturais não são transcritos nem subsequentemente traduzidos, fazendo com que o triptofano deixe de ser produzido pela célula.

Você deve estar imaginando qual seria a importância do mecanismo de atenuação no óperon trp se a repressão no início da transcrição poderia causar um efeito similar sobre a expressão gênica. Na realidade, a repressão dos genes estruturais nunca é completa, de modo que a atenuação funciona como um mecanismo adicional que aumenta a eficiência do controle da regulação gênica. Além disso, a repressão responde aos níveis celulares diferenciados de triptofano, enquanto a atenuação é modulada pela quantidade de $tRNA^{trp}$ disponíveis. Em algumas situações, a resposta em um desses níveis pode ser mais proveitosa para a célula do que em outro. Por último, enquanto um mesmo repressor pode atuar em vários óperons, o óperon trp é regulado especificamente por apenas um atenuador.

Regulação bacteriana em nível traducional

A maior parte da regulação da expressão gênica em procariotos acontece em nível de transcrição. No entanto, da mesma maneira que a transcrição pode ser regulada em seu início e prosseguimento, a tradução também pode ser alvo da regulação da expressão gênica em bactérias, mesmo que em uma menor escala. Esse outro nível de *regulação pós-transcricional* atua como um mecanismo de ajuste fino da regulação e pode ser percebido quando analisamos, por exemplo, o óperon Lac. Vimos que os genes estruturais presentes neste óperon codificam 3 enzimas distintas: β-galactosidase, permease e transacetilase. Se considerarmos que esses genes estruturais têm o mesmo promotor e são transcritos juntos, seria de se esperar que as quantidades dessas 3 enzimas fossem iguais na célula. Mas, não é isso que ocorre. Em um meio rico em lactose, podemos encontrar cerca de 3.000 moléculas de β-galactosidase, 1.500 de permease e 600 de transacetilase. Isso demonstra claramente que, após a transcrição, os transcritos de mRNA podem sofrer outras formas de regulação, que funcionam como um ajuste mais preciso do controle da expressão gênica.

Um mesmo transcrito de mRNA gerado a partir dos genes estruturais ordenados lado a lado pode provocar quantidades de produtos gênicos individuais por diferentes mecanismos, que incluem a *degradação diferenciada* de segmentos do mRNA poligênico, a *eficiência desigual do início da tradução* dos diferentes segmentos de mRNA e a *velocidade* distinta na qual o ribossomo se movimenta pelo mRNA poligênico em virtude da presença de diferentes estruturas secundárias.

RNA antissenso
RNA que tem sequência complementar a um RNA senso, tendo a capacidade de se ligar a ele, regulando sua expressão

Um modo de regular a expressão gênica em nível pós-transcricional está relacionado com a atuação de moléculas de RNA antissenso. Essas moléculas unifilamentares apresentam uma sequência de nucleotídios que é complementar ao transcrito de mRNA produzido por um gene-alvo. Dessa maneira, as moléculas de RNA antissenso têm a capacidade de se ligar ao mRNA-alvo *antes* do início da tradução, impedindo que o ribossomo se acople no mRNA-alvo para traduzi-lo. Para exemplificar esse mecanismo regulatório por RNA antissenso, vamos considerar o controle da expressão do gene ompF em *E. coli*. O gene ompF codifica proteínas de membrana que atuam como poros de difusão, capazes de regular a capacidade de a *E. coli* se adaptar à osmolaridade externa. Normalmente, quando a osmolaridade está baixa, o gene ompF está sendo transcrito e traduzido, mas caso a osmolaridade externa aumente, um outro gene chamado de micF é ativado e transcrito paralelamente ao gene ompF. O mRNA produzido pelo gene micF apresenta uma sequência de nucleotídios complementar àquela presente na região 5′UTR do transcrito de mRNA do gene ompF. Ao se ligar a essa região, o transcrito micF impede que o ribossomo se acople ao sítio de ligação no mRNA do gene ompF, impossibilitando o início da tradução. Como consequência, uma menor quantidade de proteína ompF será sintetizada, diminuindo o movimento de substâncias pela membrana.

Regulação bacteriana em nível pós-traducional

Além dos mecanismos de controle da expressão gênica descritos, as bactérias também podem refinar ainda mais a regulação da expressão de seus genes por um *controle pós-traducional*. Nessa situação, o mecanismo mais comum é a *inibição por feedback*, que acontece quando o produto final de uma via biossintética é capaz de inibir a atuação da primeira enzima participante da sua via, impedindo que mais produto final seja produzido. A biossíntese do triptofano pode exemplificar esse mecanismo de regulação. O triptofano tem a habilidade de se ligar à enzima antranilato sintetase, a primeira a participar de sua via biossintética. Com isso, a enzima sofre mudança alostérica e perde a atividade, bloqueando rapidamente a síntese de mais triptofano. No mecanismo de inibição por *feedback*, a enzima inativa pode voltar a ter atividade a partir de uma nova mudança alostérica causada pela ligação da enzima ao substrato ou a outra pequena molécula efetora, aumentando a síntese do produto final da via biossintética.

Regulação da expressão gênica em eucariotos

Em bactérias, vimos que a regulação da expressão gênica está intimamente ligada à necessidade imposta pelas condições do ambiente. Em organismos multicelulares, uma complexidade maior da regulação é dada não somente pela necessidade de especialização dos diferentes tecidos, como também pela expressão temporalmente diferenciada dos genes de eucariotos. Embora todos os tecidos sejam originados a partir de uma única célula (zigoto) e apresentem a mesma informação biológica, as células especializadas têm formas e funções completamente distintas nos diferentes tecidos. Essa diferenciação é uma consequência direta da regulação dos genes que estão sendo expressos em cada tipo celular, sendo chamada de *regulação espacial*. Por exemplo, um neurônio é diferenciado em forma e função de uma célula cardíaca porque esses dois tipos celulares expressam subgrupos de genes diferentes. Isso demonstra que a simples presença de um gene em uma célula não é um compromisso de que ele sempre será obrigatoriamente expresso. Além disso, a maioria dos eucariotos tem a capacidade de ligar e desligar seus genes em resposta às exigências peculiares de cada estágio do ciclo de vida, o que é denominado de *regulação temporal*, exemplificado pelos genes da globina. Nesse sentido, existem genes em eucariotos que já estiveram ligados e agora não funcionam mais; outros que estão ativos no presente momento; genes que estão desligados, mas ainda funcionarão em determinada fase do desenvolvimento e, por último, genes que nunca funcionaram e nunca funcionarão naquele determinado tipo celular.

Para aumentar ainda mais a complexidade na regulação da expressão gênica em eucariotos, alguns genes podem ser ativados ou silenciados em resposta a *sinais ambientais*. Certos genes em plantas, por exemplo, somente estarão ativos na presença de luz, outros apenas na presença de altas temperaturas. Não podemos esquecer ainda dos fatores de crescimento, hormônios e

demais moléculas reguladoras que são produzidos em um tecido e podem modular a expressão gênica de células localizadas a distância.

A complexidade da regulação gênica em eucariotos faz com que ela seja menos compreendida que o controle gênico em bactérias e algumas diferenças importantes podem ser reconhecidas entre os dois sistemas. Primeiramente, os eucariotos apresentam poucos genes estruturais organizados em óperons, de modo que a maioria de seus genes apresenta promotores individuais, sendo transcritos de maneira independente. Além disso, a transcrição e a tradução em eucariotos ocorrem majoritariamente em espaços físicos distintos, separados pela membrana nuclear. Com isso, a regulação gênica nesses organismos não pode interferir simultaneamente nos dois níveis, como vimos em bactérias. Outra peculiaridade em eucariotos é que os sinais ambientais devem passar por uma série de barreiras celulares até chegarem aos cromossomos em organismos multicelulares. Por último, a regulação da expressão gênica em eucariotos pode ter início em um nível anterior ao que acontece na maioria dos genes bacterianos, que é representado pela remodelagem dinâmica da estrutura da cromatina.

Na Tabela 2.1, podemos ter uma visão geral dos diferentes níveis em que a regulação da expressão gênica acontece em eucariotos.

■ **Tabela 2.1** Diferentes níveis de controle da expressão gênica evidenciados em eucariotos

Nível de regulação gênica	Mecanismo de regulação
Cromatina	Regulação epigenética mediada principalmente por modificações químicas nas caudas de histonas e metilação do DNA, capazes de compactar/descompactar a cromatina, tornando-a menos ou mais acessível à maquinaria transcricional. Proteínas Policomb e Tritorax
Transcrição	Ligação de fatores transatuantes a sequências cisatuantes Uso de promotores alternativos
Pós-transcrição	Encadeamento alternativo dos éxons Degradação do mRNA Utilização de sítios de poliadenilação alternativos Edição do mRNA
Tradução	RNA antissenso, RNA de interferência
Pós-tradução	Modificações químicas pós-traducionais Clivagem proteolítica do produto primário Degradação da proteína

Cromatina e a regulação epigenética da expressão gênica

Durante o ciclo de vida de um organismo vivo, fatores epigenéticos provocam uma maneira herdável, reversível e dinâmica de modular a expressão gênica sem que haja modificação na sequência do DNA. Como vimos no Capítulo 1, no núcleo celular, a molécula de DNA interage com octâmeros de histonas, formando os *nucleossomos* que, por conseguinte, se compactam de maneira mais organizada ainda para compor a *cromatina*. Nesse sentido, o primeiro nível do controle da expressão em eucariotos envolve o *posicionamento dinâmico das histonas*, organizando o genoma em uma cromatina mais aberta ou mais fechada. Essa flexibilidade governa a atividade transcricional, visto que quando a cromatina está condensada, o acesso da maquinaria transcricional (RNA polimerase e fatores transcricionais) aos sítios de ligação na molécula de DNA se torna prejudicado, inibindo a expressão do(s) gene(s) ali presente(s).

Epigenética
Estudo das modificações hereditárias na regulação de um gene que não são atribuíveis a mudanças em sua sequência de DNA

Sendo assim, para que a transcrição aconteça, a cromatina obrigatoriamente deve estar em estado descompactado. Caso contrário, a transcrição é reprimida.

Regiões de eucromatina podem estar transcricionalmente ativas ou não, dependendo dos mecanismos de controle da expressão gênica aos quais estão submetidas, ao passo que regiões de heterocromatina estarão obrigatoriamente inativas.

Uma das maneiras de garantir a flexibilidade da cromatina de acordo com as necessidades celulares é representada pela capacidade que as caudas aminoterminais das histonas apre-

sentam de sofrerem várias modificações pós-traducionais covalentes, embora dinâmicas. A multiplicidade de combinações envolvendo diferentes modificações pós-traducionais nessas proteínas acaba por gerar um *código de histonas*. Dentre as modificações mais conhecidas, podemos citar a adição/remoção de grupamentos acetil, fosfato e metil nas caudas das histonas em aminoácidos específicos.

A *acetilação* de resíduos de lisina presentes nas caudas das histonas é um marco de regiões transcricionalmente ativas, enquanto as histonas hipoacetiladas são encontradas na eucromatina inativa ou em regiões de heterocromatina. Como as histonas acetiladas têm uma menor afinidade pelo DNA do que, possivelmente, entre elas, a cromatina pode ser capaz de adotar uma estrutura mais relaxada, mais adequada à transcrição. Por outro lado, a desacetilação de histonas promove a repressão da expressão gênica em virtude de a cromatina tornar-se mais condensada em decorrência da livre interação das moléculas de histonas com a molécula de DNA. O equilíbrio entre a acetilação e desacetilação de histonas é governado pelas ações contrárias de enzimas histonas acetiltransferases, que adicionam os grupamentos acetil às histonas e de desacetilases de histonas, capazes de remover os grupamentos acetil das caudas das histonas.

As histonas apresentam carga elétrica positiva, o que possibilita a sua interação com os grupos fosfato de carga negativa presentes no arcabouço da molécula de DNA. De modo semelhante à acetilação, a *adição de grupamentos fosfato* carregados negativamente às caudas das histonas pode neutralizar a carga destas últimas. Por conseguinte, as histonas fosforiladas perdem a afinidade pelo DNA e a cromatina se torna mais descompactada, possibilitando o acesso da maquinaria transcricional. Nesse contexto, a fosforilação da serina 10 na histona H3 tem sido correlacionada com ativação gênica em células de mamíferos.

Contrariamente à acetilação e à fosforilação, a *adição de grupamentos metil* às caudas de histonas pode representar tanto um sinal para ativação transcricional quanto para inativação gênica, dependendo dos aminoácidos que são metilados na cauda de histonas e do grau de metilação presente. A metilação da lisina 9 na extremidade aminoterminal da histona H3 é um marco para o silenciamento do DNA e está globalmente distribuída ao longo de regiões de heterocromatina como os centrômeros e telômeros, além de estar presente em promotores silenciados e no cromossomo X inativo. Em contraste, a metilação da lisina 4 da histona H3 provoca atividade gênica e pode ser encontrada predominantemente nos promotores de genes ativos. Uma possível explicação para esta dupla função da metilação seria a de que a metilação da lisina 9 da histona H3 é mediada pela ligação da proteína de heterocromatina 1 (HP1), enquanto a da lisina 4 da histona H3 aparentemente não o é. Além disso, os resíduos de lisina presentes nas caudas de histonas podem ser mono, di ou trimetilados e essa diferença no padrão de metilação também pode influenciar o controle da expressão gênica. Um número crescente de enzimas capazes de adicionar grupamentos metil (metiltransferases de histonas) e de removê-los (desmetilases de histonas) tem sido descrito.

Além das modificações químicas citadas, a compactação/descompactação da estrutura da cromatina pode ser regida por *complexos de remodelagem de cromatina SWI/SNF ATP-dependentes*, os quais consistem em complexos de multisubunidades capazes de modificar a topologia dos nucleossomos, alterando, assim, a interação do DNA com as histonas. Dessa forma, esses complexos de remodelagem da cromatina se ligam a sítios particulares da molécula de DNA e deslocam temporariamente os nucleossomos ali existentes na presença de ATP, possibilitando que os fatores de transcrição tenham acesso aos seus promotores-alvo dentro da cromatina. Com isso, os níveis de transcrição podem ser aumentados.

Metilação do DNA

Além das modificações de histonas, os genomas de todos os vertebrados também podem ser modificados pela adição de um *grupamento metil* (CH3) ao carbono 5 da citosina. A metilação das citosinas ocorre predominantemente nas citosinas presentes em regiões ricas em dinucleotídios CG ligados covalentemente lado a lado, chamadas de ilhas CpG. As ilhas CpG estão presentes em mais de metade das regiões promotoras dos genes que compõem nosso genoma. Sequências de DNA que estão transcrevendo ativamente necessitam estar desmetiladas ao menos nessas regiões promotoras. Sendo assim, quanto maior for a extensão de metilação

Ilhas CpG
Segmento da molécula de DNA que é rico em dinucleotídios CG ligados covalentemente e é frequentemente encontrado nas regiões promotoras dos genes

de ilhas CpG, menor será a taxa de transcrição dos genes, cujos promotores estão metilados. Quanto menor a taxa de metilação, maior a possibilidade de haver transcrição.

A doação de carbonos para a manutenção dos padrões de metilação é primariamente realizada pelo metabolismo do *folato*, o qual envolve a ação integrada de vários produtos gênicos e de outros micronutrientes como a vitamina B12, vitamina B6, colina e metionina, obtidos pela dieta. Todos esses elementos são direta ou indiretamente necessários para a conversão da homocisteína a metionina, que é o precursor imediato de S-adenosil metionina (SAM), o maior doador intracelular de radicais metil para a manutenção da metilação do DNA, proteína e lipídios em nosso organismo. Após doar um grupamento metil, o SAM é convertido em S-adenosil-homocisteína (SAH), que é, então, transformado em homocisteína.

Os radicais metil são adicionados às citosinas por 3 enzimas DNA metiltransferases principais (Dnmt1, Dnmt3a e Dnmt3b) e 2 tipos de metilação podem ser reconhecidos: a *metilação de manutenção* e a *metilação de novo*. A replicação do DNA cria moléculas filhas hemimetiladas de moléculas molde completamente metiladas. A restauração dos padrões de metilação é efetuada pela ação da metilação de manutenção. Os sítios metilados no DNA parental funcionam como molde para a metilação correta do novo filamento, o que garante que o padrão da metilação anterior seja mantido corretamente. Como resultado, ambos os filamentos-filhos são metilados nos mesmos sítios que o DNA parental. A enzima responsável pela metilação de manutenção é a Dnmt 1. As Dnmt3a e Dnmt3b são expressas em células embrionárias e têm a função de adicionar grupos metil em novas posições em ambos os filamentos de DNA após a desmetilação global que ocorre no período pré-implantação, processo conhecido como metilação *de novo*.

Methyl CpG binding proteins (MBP)
Grupo de proteínas capazes de se ligar a ilhas CpG e reprimir a transcrição gênica de maneira direta ou indireta

Contudo, como a metilação leva à repressão transcricional? A repressão das ilhas CpG metiladas localizadas nas sequências promotoras é mediada por proteínas de ligação às ilhas CpG metiladas (MBPs, do inglês *methyl CpG binding proteins*). As MBPs têm a habilidade de se ligar ao DNA metilado e bloquear a expressão dos genes por elas regulados de duas maneiras. Em primeiro lugar, a simples ligação das MBPs ao DNA pode perturbar espacialmente a interação dos fatores transcricionais com a cromatina, levando à inativação do gene-alvo ali presente. Além disso, após se ligarem a regiões metiladas específicas, as MBPs têm o poder de recrutar fatores de compactação da cromatina, como, por exemplo, desacetilases de histonas. A compactação subsequente da cromatina impede o acesso dos fatores transcricionais, favorecendo a repressão do gene-alvo.

Imprinting parental
Determinação da expressão de um gene pela origem parental

A metilação do DNA tem papel importante em vários processos biológicos, incluindo o controle da expressão gênica espacial e temporal, a inativação de um dos cromossomos X na mulher, a segregação centromérica adequada durante a divisão celular, a proteção contra elementos genéticos móveis, o *imprinting* parental e o envelhecimento. Em virtude da participação da metilação do DNA em todos esses processos, não é de se estranhar que padrões alterados de metilação venham sendo associados a um número crescente de doenças, dentre eles o câncer. Entre os processos biológicos, nos quais a metilação do DNA participa, escolhemos dois para descrevermos mais detalhadamente: a *inativação do cromossomo X* e o *imprinting parental*.

Inativação do cromossomo X

Conforme vimos no Capítulo 4, ao longo do processo evolutivo, o cromossomo Y sofreu uma redução de tamanho atrelada à presença de mutações que o impediram de recombinar meioticamente com grande parte do cromossomo X. Nesse processo, apenas as extremidades de ambos os cromossomos mantiveram a capacidade de recombinar entre si durante a gametogênese. Para que haja *compensação de dose* na expressão de genes dos cromossomos sexuais entre os sexos masculino e feminino, um dos cromossomos X na mulher necessita ser desligado. Por meio da inativação do X, a mulher passa a ser funcionalmente hemizigota em nível celular para genes localizados no cromossomo X, assim como o homem o é. Com isso, o quantitativo de genes expressos a partir dos cromossomos sexuais em ambos os sexos passa a ser equivalente. Essa inativação ocorre por um processo que conjuga vários mecanismos epigenéticos de silenciamento gênico, dentre os quais está a hipermetilação de um dos X, ainda no início do desenvolvimento embrionário.

Hipótese de Lyon
Hipótese formulada por Mary Lyon em 1961 que preconiza que um cromossomo X em cada célula feminina é desligado de modo aleatório para garantir uma compensação de dose entre os sexos feminino e masculino

Epistasia
Interação gênica em que o efeito de um gene é suprimido pela ação de outro gene, não alélico

Os preceitos que regem a inativação de um dos cromossomos X na mulher começaram a ser elucidados por Mary Lyon em 1961 que formulou a Hipótese de Lyon. O cromossomo X inativo pode ser tanto o de origem materna como o de origem paterna e a escolha por qual deles terá seus genes desligados em uma célula é *aleatória*. Isso quer dizer que essa inativação randômica faz com que toda mulher represente um *mosaico* com relação à expressão de genes do cromossomo X, pois terá globalmente em seu organismo 50% de células que expressam o cromossomo X paterno e 50% de células que expressam o cromossomo X maternalmente herdado. Desvios nesse equilíbrio entre a porcentagem de células com o X materno e o X paterno inativos são comuns em distúrbios envolvendo o cromossomo X.

A partir do momento que o cromossomo X de uma determinada linhagem parental é inativado em uma célula-mãe, todas as células-filhas descendentes dessa célula terão o mesmo cromossomo parental inativo. Ou seja, se em dada célula durante o desenvolvimento embrionário o X inativo foi o de origem paterna, por exemplo, essa célula gerará novas células que também terão o cromossomo X paterno inativo.

O padrão de inativação do cromossomo X em forma de mosaico pode ser representado pela distribuição de manchas na pelagem de gatos domésticos. Nesses gatos, um dos genes que contribuem para a cor da pelagem está localizado no cromossomo X e determina a cor marrom-preta (dominante) ou amarelo-laranja (recessiva). A fêmea heterozigota para esse gene terá áreas com ambas as cores, ocasionando uma pelagem malhada. O padrão de distribuição das manchas será reflexo das diferentes células que expressam um dos dois alelos diferentes, dependendo de qual cromossomo X está inativo. Manchas brancas também podem ocorrer em virtude da epistasia de um gene autossômico que inativa a síntese de pigmento. Nos gatos do sexo masculino, todas as manchas existentes apresentarão a cor determinada pelo alelo presente no seu único cromossomo X. Nas fêmeas, a coloração dependerá tanto da composição alélica, como do mecanismo de inativação do cromossomo X.

Entretanto, de que maneira a inativação do cromossomo X acontece? A inativação ocorre em um *estágio inicial do desenvolvimento* (entre o 13º e 16º dias do desenvolvimento embrionário) e é governada pelo *centro de inativação do X* (Xic), região localizada no cromossomo X em Xq13. O Xic é um *locus* cis atuante que contém a informação necessária para contar quantos cromossomos X estão presentes na célula e inativar todas as cópias, com exceção de uma. Por esse motivo, nos indivíduos com número anormal de cromossomos X (45,X; 47,XXX; 47,XXY etc.), um único X permanece ativo, minimizando as consequências fenotípicas. A inativação do X começa na região Xic e se propaga para todo o cromossomo X. Após a fertilização, a expressão no X inativo de um dos genes contidos na região Xic, o gene XIST (do inglês *X-inactive specific transcript)*, é requerida para iniciar o silenciamento gênico de um dos cromossomos X presentes no sexo feminino. Esse gene expressa um *longo RNA não codificante* que reveste o cromossomo X que será desligado, induzindo sua inativação, o que sugere que ele tenha papel estrutural. Paradoxalmente, o gene XIST está ligado no X inativo e desligado no X ativo. Dessa maneira, enquanto o cromossomo X inativo se distingue pela expressão do RNA Xist, o X ativo se caracteriza por expressar um longo transcrito antissenso ao Xist, chamado de Tsix. A expressão de Tsix é requerida para o desligamento do XIST e evita a inativação do cromossomo X ativo.

As características do cromossomo X inativo incluem *hipermetilação das ilhas CpG* e *hipoacetilação das histonas H4*, presumidamente ocorrendo mais tardiamente como parte do mecanismo de inativação após o silenciamento de genes mediado pelo Xist. Portanto, acredita-se que a metilação do DNA no cromossomo X inativo tenha importância fundamental na manutenção do padrão de inativação gênica. Além disso, o cromossomo X reúne vários mecanismos adicionais para silenciar a cromatina, que incluem a expressão de moléculas de RNA não codificante, uma sinalização para replicação tardia, demais modificações de histonas e o recrutamento da variante de histona macroH2A. Hoje, sabe-se, no entanto, que cerca de 1/6 dos genes do cromossomo X inativo escapam à inativação, dentre eles o gene XIST. Esses genes que não se distribuem aleatoriamente no cromossomo X teriam, portanto, o dobro da expressão nas mulheres do que nos homens. Um percentual dos genes que escapam à inativação do X está situado fora da região pseudoautossômica e não apresentam cópias correspondentes no cromossomo Y. Isso faz com eles tenham maior expressão em mulheres, o que pode

causar diferenças fenotípicas entre os dois sexos. Em compensação, enquanto alguns genes do cromossomo Y têm expressão restrita aos testículos, outros se expressam amplamente no organismo. Essas peculiaridades dos cromossomos X e Y podem ser responsáveis por diferenças comportamentais entre os dois sexos, podendo ter relação até mesmo com uma suscetibilidade diferenciada de homens e mulheres em relação a algumas doenças. A presença de genes ativos no cromossomo X inativo também explicaria o fato de um indivíduo com a síndrome de Klinefelter (47,XXY) diferir fenotipicamente de um homem com cariótipo 46,XY e, de uma mulher com síndrome de Turner (45,X), na qual o único cromossomo X presente está ativo, diferir fenotipicamente de uma mulher tipicamente 46,XX.

Enquanto em mamíferos placentários a compensação de dose entre os dois sexos é obtida pela inativação de um dos cromossomos X na fêmea, em outros organismos a compensação de dose pode ser alcançada por outros mecanismos. Nas moscas de frutas, por exemplo, a compensação de dose ocorre pela *duplicação* da atividade dos genes do cromossomo X nos machos. Já no verme *Caenorhabditis elegans*, ambos os cromossomos X na fêmea têm *metade* da atividade de seus genes para efetuar a compensação de dose entre os dois sexos.

Imprinting parental e dissomia uniparental

Tanto o genoma materno quanto o paterno são imprescindíveis para o desenvolvimento de um organismo, tanto em nível embrionário como pós-natal. Talvez por isso tenhamos o hábito de considerarmos que ambos os alelos (materno e paterno) de um gene bialélico se expressam a todo o momento, a não ser que uma ou ambas as cópias desse gene tenham sido alvo de mutações que levem ao silenciamento gênico. Contudo, sabe-se que para certos genes, em determinados períodos do desenvolvimento ou em alguns tecidos ou até mesmo em todas as células, é necessário que apenas um dos alelos parentais se expresse. Nessas situações, a expressão de um dos alelos é reprimida e, portanto, apenas metade do produto do gene será expresso.

Em alguns casos, a escolha por qual alelo será desligado (materno ou paterno) não é aleatória, mesmo que as sequências de ambos os alelos parentais sejam idênticas.

Sendo assim, o alelo, cuja expressão é reprimida, deve ser obrigatoriamente o herdado patrilinearmente ou o herdado matrilinearmente, demonstrando o efeito do genitor sobre a expressão dos genes. A esse mecanismo é dado o nome de *imprinting* parental. Então, além da necessidade de termos cromossomos com número e estrutura corretos, também é essencial que eles tenham a origem parental adequada. Um exemplo do efeito do genitor na expressão gênica pode ser demonstrado na síndrome de Turner (45,X). Quando o único cromossomo X é herdado do pai, o portador da síndrome apresenta maiores habilidades verbais e comportamentais se comparado aos casos, nos quais o cromossomo X foi herdado da mãe.

Como o *imprinting* parental é estabelecido? A marca parental que é concedida a um gene é representada pela *metilação do DNA* de um ou mais dinucleotídios CpG na proximidade do gene que é regulado por este mecanismo. O processo de *imprinting* é regulado por regiões chamadas de *regiões de controle de imprinting* e acontece na linhagem germinativa parental. Nesse momento, determinado gene que é submetido ao *imprinting* será metilado apenas na linhagem germinativa paterna ou na materna, nunca em ambas. Sendo assim, durante a fecundação, os gametas apresentam-se diferencialmente metilados, de modo que o embrião, em um estágio inicial do desenvolvimento apresenta diferenças de metilação nos dois alelos (materno e paterno). O espermatozoide e o óvulo apresentam, portanto, padrões específicos e diferentes de metilação. Já na fase de mórula e blástula inicial, ocorre uma ampla *desmetilação* do genoma e as células germinativas primordiais que irão originar os gametas permanecem desmetiladas até a diferenciação das gônadas quando ocorrerá *metilação de novo*. Com isso, um gene metilado herdado do pai pode adquirir um novo padrão de metilação quando é transmitido para uma prole de sexo oposto, ou seja, uma mulher que carrega em suas células germinativas uma marca epigenética feminina pode ter filhos do sexo masculino que originarão seus gametas com a marca epigenética masculina. Isso nos mostra a importância de apagar e restabelecer o *imprinting* parental a cada nova geração.

O *imprinting* parental é o principal motivo pelo qual a partenogênese é inviável. Além disso, devemos lembrar que em casos de clonagem animal ainda não é possível gerenciar, pelo menos

de maneira fina, os padrões de *imprinting* parental, o que explicaria o fato de experimentos de clonagem animal apresentarem taxa de sucesso ainda considerada baixa. Uma teoria semelhante é sugerida para embriões oriundos de técnicas de reprodução assistida. Alguns estudos epidemiológicos sugerem que tais procedimentos poderiam perturbar o estabelecimento dos padrões epigenéticos da expressão gênica no embrião. Com isso, as crianças concebidas *in vitro* apresentariam maior risco relativo de apresentar baixo peso ao nascimento, defeitos congênitos e doenças relacionadas com a falha no *imprinting* parental, apontando que alterações no padrão epigenético poderiam estar associadas à reprodução assistida. No entanto, essa especulação ainda necessita ser melhor investigada.

Padrões alterados de *imprinting* parental já foram associados a várias doenças genéticas. O exemplo mais característico deste evento é representado pelas *síndromes de Prader-Willi* e *Angelman*. A síndrome de Prader-Willi é caracterizada por deficiência intelectual, compulsão alimentar, obesidade, mãos e pés pequenos, baixa estatura e atraso no desenvolvimento sexual, enquanto a síndrome de Angelman é representada por uma grave deficiência intelectual, riso frequente e convulsões. Embora esses quadros fenotípicos sejam distintos, as duas síndromes surgem majoritariamente de uma deleção em uma mesma região do braço longo cromossomo 15 que é submetida à *imprinting* parental. Quando a deleção ocorre no cromossomo 15 paterno, o indivíduo só expressará o alelo materno e manifestará a síndrome de Prader-Willi. Quando a perda envolve o cromossomo 15 materno, o indivíduo expressará apenas o alelo paterno e apresentará a síndrome de Angelman.

Dissomia uniparental
Erro que leva uma célula ou organismo com número cromossômico normal a ter uma mesma origem parental incorreta para um par de cromossomos em particular.

Isodissomia
Evento de dissomia parental, no qual os dois alelos ou cromossomos com a mesma origem parental são iguais

Heterodissomia
Evento de dissomia parental, no qual os dois alelos ou cromossomos com a mesma origem parental são diferentes

As duas síndromes também podem surgir por uma não disjunção do cromossomo 15 em ambos os gametas que se unem para formar um zigoto. Esse erro pode fazer com que um par de cromossomos 15 homólogos ou parte deles seja originado apenas de um dos genitores, o que é denominado dissomia uniparental. Caso um mesmo cromossomo do genitor esteja em duplicata, chamamos o evento de isodissomia. Caso ambos os cromossomos homólogos do genitor estejam presentes, o evento é denominado heterodissomia. A dissomia uniparental é uma situação rara, pois requer que aconteçam dois eventos incomuns simultaneamente, um deles levando à perda de um cromossomo ou segmento cromossômico e outro levando à adição desse mesmo material genético que foi afetado no primeiro evento. Sendo assim, podemos observar a dissomia uniparental em apenas uma parcela (20% a 30%) dos casos de síndrome de Prader-Willi e Angelman, assim como em outras doenças relacionadas com falhas no *imprinting*. Caso haja perda da expressão do alelo que seria de origem paterna, o indivíduo manifestará a síndrome de Prader-Willi. Caso ocorra o inverso, ou seja, o indivíduo tenha perdido a contribuição materna dessa região imprintada no cromossomo 15, a manifestação fenotípica será da síndrome de Angelman. Outra maneira de originar as duas síndromes é mediante a presença de falhas na metilação do DNA durante o estabelecimento da marca epigenética do cromossomo 15 na linhagem germinativa de um dos genitores.

Outras formas de regulação epigenética

Proteínas Policomb e Tritorax

Uma classe adicional de regulação epigenética envolve a ação de proteínas dos grupos *Policomb* e *Tritorax*. Essas proteínas modificadoras da cromatina fazem parte de um *sistema de memória celular* responsável por controlar a acessibilidade da cromatina e a manutenção da transcrição nos primeiros estágios da vida embrionária, durante o desenvolvimento e na vida adulta. Para esse propósito, essas proteínas têm efeitos *antagonistas* em genes-alvo. Enquanto as proteínas Policomb atuam como *repressoras* estáveis, que perpetuam o estado de repressão durante as divisões celulares, as proteínas Tritorax promovem a manutenção da *atividade gênica*. Evidências demonstram que componentes do complexo Policomb e polipeptídios do grupo Tritorax são metiltransferases de histonas, sugerindo que o sistema de memória utiliza um caminho similar para regular a ligação de ambos os grupos de proteínas nos seus sítios-alvo.

Controle da expressão gênica de longo alcance

A maioria dos genes eucarióticos parece estar situada em regiões de eucromatina. Estudos nos quais genes eucromáticos são reposicionados em um ambiente de heterocromatina por

manipulação genética em organismos modelos ou através de translocações espontâneas demonstram que eles passam a funcionar de maneira anormal ou deixam de funcionar completamente, mesmo se toda sua extensão, incluindo suas sequências flanqueadoras, tenha permanecido intacta. Isso sugere que a estrutura da cromatina pode exercer um *controle de longo alcance* sobre a expressão gênica e que os cromossomos estão organizados em domínios funcionais da expressão gênica, os *domínios de cromatina*. Quando um gene eucromático é reposicionado em regiões próximas de centrômeros, telômeros ou blocos de heterocromatina, por exemplo, sua expressão pode ser abolida pela alteração da estrutura de um grande domínio de cromatina. Sendo assim, a expressão do gene eucromático pode ser suprimida por um efeito de posição de longo alcance, como consequência do novo ambiente cromossômico criado pela translocação.

Regulação da expressão gênica em eucariotos em nível transcricional

Agora que já vimos como os eucariotos podem regular a expressão de seus genes em nível de estrutura da cromatina, descreveremos outras formas de regulação em nível transcricional. Como em procariotos, a maior parte do controle da expressão gênica em eucariotos acontece em nível de *transcrição*. O início da transcrição eucariótica requer que fatores proteicos transcricionais e a RNA polimerase formem um aparato de transcrição basal para dar início à transcrição. A regulação da expressão gênica nessa etapa está associada diretamente à montagem ou estabilidade desse aparato. Normalmente, o aparato de transcrição basal favorece a transcrição de níveis mínimos de RNA, e proteínas transcricionais ativadoras são necessárias para que a transcrição ocorra em níveis diferentes dos basais. Como já visto, essas proteínas reguladoras (*fatores de transcrição transatuantes*) são capazes de modular os níveis de transcrição ao se ligarem em regiões específicas do DNA (*elementos cisatuantes*) localizadas nas proximidades de um gene ou até mesmo em seu interior.

A interação entre os fatores proteicos transatuantes e sequências cisatuantes de ácidos nucleicos também é a base comum para a regulação da expressão gênica em outras instâncias que não o início da transcrição, incluindo o processamento do mRNA, o transporte do mRNA e a tradução. Os fatores transatuantes em eucariotos apresentam segmentos, constituídos por cerca de 60 a 90 aminoácidos, chamados de *domínios* que têm a função de se ligar ao DNA. Dentro de um domínio de ligação, encontramos uma ou mais estruturas simples resultantes de associações entre aminoácidos da cadeia polipeptídica denominadas *motivos*, que garantem uma interação forte entre a proteína reguladora e o DNA. Esses motivos utilizam a α-hélice ou ocasionalmente a folha β pregueada para se ligarem à curvatura maior da molécula de DNA e têm especificidades distintas para possibilitar a ligação da proteína a seus genes-alvo.

Entre os motivos estruturais mais conhecidos, destacam-se o "dedo de zinco", o hélice-volta-hélice, o zíper de leucina e o hélice-alça-hélice (Figura 2.23). O motivo "dedo de zinco" (*zinc finger*, em inglês) é representado por uma alça polipeptídica curta formada quando duas cisteínas em uma parte do polipeptídio e duas histidinas em outro segmento do polipeptídio se unem a um íon de zinco, formando uma estrutura que se assemelha a um dedo. Já o motivo hélice-giro-hélice é composto por duas α-hélices separadas por uma sequência curta de aminoácidos que induz um giro agudo, dando o nome ao motivo. Este motivo é comumente encontrado em vários fatores transcricionais e em genes chamados de *homeoboxes*, responsáveis por regular o desenvolvimento (morfogênese) de animais, fungos e plantas. O motivo zíper de leucina, por sua vez, é constituído por um trecho de aminoácidos com uma leucina em cada sétima posição, enquanto o motivo estrutural hélice-alça-hélice apresenta duas regiões helicoidais de aminoácidos separadas por uma alça não helicoidal.

Os elementos cisatuantes podem ser os próprios *promotores* (centrais e não centrais), *reforçadores*, *silenciadores*, *insuladores* e *elementos de resposta*. Os *reforçadores* são elementos *reguladores positivos* que, ao interagirem com fatores proteicos específicos, aumentam os níveis de transcrição basal realizado pelos promotores centrais. Esses elementos reforçadores podem estar localizados até mesmo a distância dos genes que são por eles regulados e podem conferir especificidade de tecido e estágio de desenvolvimento.

Entretanto, como os elementos reforçadores poderiam exercer controle sobre a transcrição iniciada no promotor? Algumas evidências apontam que as proteínas reforçadoras que se li-

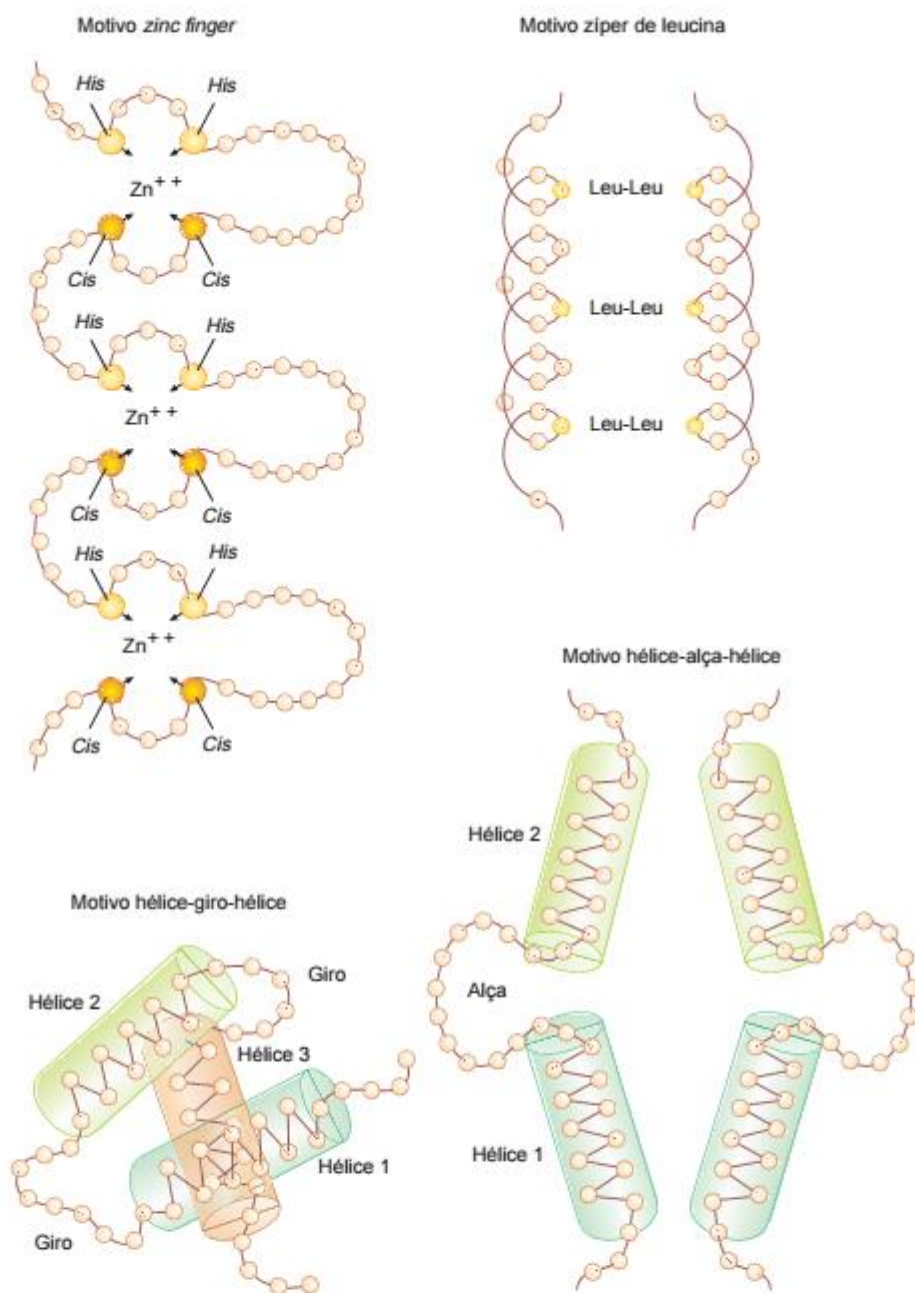



Figura 2.23 Principais motivos encontrados em proteínas de ligação ao DNA.

gam aos elementos reforçadores podem influenciar a atividade de polipeptídios que se ligam aos promotores, incluindo a RNA polimerase e os fatores de transcrição basal. Essa interação parece ser possível pelo *dobramento do DNA*, que possibilitaria o contato físico entre esses dois grupos de fatores proteicos. Algumas vezes, os fatores de transcrição reforçadores também podem atuar de modo indireto no aparato basal pela participação de proteínas *correforçadoras*. Essas proteínas provocam uma conexão entre os fatores transcricionais reforçadores e o aparato basal, mesmo que por si só não se liguem ao DNA. Algumas proteínas reforçadoras, correforçadoras e fatores gerais de transcrição têm atividade de acetiltransferase de histonas, o que estimula a transcrição pela descompactação da cromatina.

Em contraposição aos reforçadores, os *silenciadores* funcionam como elementos de *controle negativo*, reduzindo os níveis de transcrição. Em genes humanos, os elementos silenciadores

podem estar localizados próximos ao promotor, a alguma distância anterior a ele e no interior de íntrons. Os silenciadores também podem proporcionar especificidade de tecido ou estágio do desenvolvimento, bloqueando a expressão de genes em tecidos, nos quais sua expressão não é requerida.

Os *insuladores*, por sua vez, são sequências do DNA que têm por objetivo bloquear a influência de agentes reforçadores e silenciadores. Desse modo, quando uma proteína insuladora é alocada entre um reforçador e um promotor, ela evita que o promotor seja ativado pelo reforçador. Isso é importante porque a maioria dos reforçadores é capaz de estimular indiscriminadamente qualquer promotor em sua vizinhança e os insuladores possibilitam aumentar a precisão da regulação gênica restrita a um promotor específico. Caso o insulador fique fora da região entre o reforçador e o promotor, ele não terá efeito sobre a expressão gênica. Da mesma maneira, quando o insulador está situado entre um gene ativo e uma região de heterocromatina, ele constrói uma barreira que protege o gene contra o efeito de inativação que se propaga a partir da heterocromatina. Alguns insuladores apresentam essas duas propriedades.

Por último, os *elementos de resposta* modulam os níveis de transcrição de acordo com um estímulo externo específico, como, por exemplo, uma mudança de temperatura, a presença de luz, concentrações extracelulares de certos íons e a presença de um hormônio ou fator de crescimento. Em geral, nesses casos, um fator de transcrição, inicialmente inativo é ativado pela rota de sinalização e se liga a elementos reguladores (elementos de resposta) nos genes-alvo, estimulando a transcrição desses genes. Um único gene eucariótico pode ser regulado por diferentes elementos de resposta e um mesmo elemento de resposta em diferentes genes possibilita que um único estímulo externo ative vários genes simultaneamente.

Proteínas de choque térmico
Polipeptídeos produzidos em muitos organismos em resposta ao calor extremo, evitando danos maiores às células

Uma classe particular de genes regulados por estímulos externos é representada pelos genes que codificam polipeptídios altamente conservados, chamados de proteínas de choque térmico ou *heat-shock*. Essas proteínas, presentes tanto em procariotos como em eucariotos, são produzidas em resposta è exposição do organismo a uma alta temperatura, estabilizando o ambiente celular. A regulação da expressão dessas proteínas ocorre em nível transcricional. Na *Drosophila*, por exemplo, uma proteína de choque térmico, a HSP70, é sintetizada quando a temperatura excede 33°C por uma família de genes hsp70. A síntese é mediada pelo *fator de transcrição heat-shock* (HSTF) que, na presença de calor, é modificado quimicamente por fosforilação e torna-se capaz de se ligar a uma sequência de nucleotídios antecedente aos genes hsp70, chamada de *elementos de resposta heat-shock*, estimulando a transcrição desses genes.

As proteínas de choque térmico são produzidas por um grande número de genes diferentes. Esses genes costumam ser expressos de modo coordenado, mesmo que não estejam aglomerados como acontece nos óperons bacterianos. Isso acontece em virtude de eles apresentarem elementos de resposta comuns em seus promotores ou acentuadores. Essas sequências regulatórias de DNA contêm sequências consenso que funcionam como sítios de ligação para os ativadores transcricionais durante os períodos de estresse ambiental, aumentando os níveis de transcrição.

Outras formas de regulação da expressão gênica em nível de transcrição/processamento do mRNA incluem:

- *Utilização diferencial de promotores alternativos*: alguns genes eucarióticos têm dois ou mais *promotores alternativos*. O uso desses promotores alternativos durante a transcrição implica na utilização de éxons diferentes, podendo dar origem a diferentes isoformas proteicas com propriedades distintas. A utilização diferencial desses promotores pode garantir especificidade de tecido ou de estágio de desenvolvimento, uma localização subcelular preferencial (nuclear ou citoplasmática), uma capacidade funcional diferenciada ou até mesmo um controle da expressão gênica sexo-específica. O gene da distrofina localizado em Xp21, por exemplo, utiliza 8 promotores alternativos
- *Recomposição alternativa dos éxons*: A maioria dos genes de eucariotos apresenta íntrons. Durante o processamento do pré-mRNA, essas regiões não codificantes são removidas e os éxons adjacentes a elas são encadeados para formar o mRNA maduro. Contudo, para alguns genes diferentes, sequências de mRNA maduro podem ser geradas durante o processamento de um transcrito de mRNA com base na *combinação diferencial dos éxons*. Os

mRNA gerados dessa recomposição alternativa, por sua vez, serão traduzidos em isoformas proteicas distintas. Essas isoformas podem apresentar diferentes localizações subcelulares, especificidade de tecido ou até mesmo funções díspares. A escolha dos éxons durante o processo de recomposição é mediada pela ligação de fatores proteicos ao pré-RNA

- *Utilização de sinais de poliadenilação alternativos*: o uso após a transcrição de dois ou mais sinais de poliadenilação alternativos localizados na região 3′UTR pode gerar mRNA diferenciados com especificidades de tecido
- *Degradação do mRNA*: Após a transcrição, os mRNA de eucariotos são exportados do núcleo até o citoplasma, onde serão traduzidos sequencialmente pelos ribossomos. A presença ou ausência de mRNA na célula determina, por conseguinte, quais proteínas serão sintetizadas naquele período. A quantidade de mRNA disponível, por sua vez, resulta do balanço entre a *síntese* e a *degradação* de mRNA. Quanto maior o tempo que um mRNA permaneça estável na célula, maior a capacidade de ele participar de rodadas subsequentes na síntese de proteínas. Isso demonstra que a estabilidade pode ter relação direta com a necessidade celular de sintetizar um polipeptídio em um dado momento. Como já comentado, o mRNA eucariótico é mais estável que o bacteriano, mas essa estabilidade é muito variável variando de alguns minutos a até dias. A degradação do mRNA depende da ação de ribonucleases e pode acontecer de diferentes maneiras. A mais comum delas envolve o encurtamento da cauda poli (A) seguido da perda do cap na região 5′ e degradação do transcrito no sentido 5′ → 3′. Outras formas de degradar o mRNA incluem: degradação iniciada na extremidade 3′ com remoção subsequente dos nucleotídios no sentido 3′ → 5′ e clivagem do mRNA em sítios internos. Além da cauda poliA e do cap, as regiões 3′UTR, 5′UTR e a própria sequência codificante também parecem interferir na estabilidade do mRNA. Alguns mRNA eucarióticos de vida curta têm, por exemplo, sequências ricas em AU na região 3′UTR que são alvo da degradação por microRNA
- *Edição diferenciada do mRNA*: mesmo que seja mais rara, a modificação diferencial da sequência de nucleotídios de um mesmo transcrito de mRNA pode levar a diferenças nas isoformas, que muitas vezes podem ser tecido específicas.

Todos os mecanismos acima citados, individualmente ou combinados, podem originar um número substancial de isoformas proteicas diferentes a partir de um único gene.

Esses mecanismos diferenciados de expressão de uma mesma sequência de DNA fazem com que tenhamos uma definição muito mais ampla do termo "gene", do que aquela na qual um gene é definido por uma sequência de nucleotídios capaz de especificar apenas uma molécula de RNA ou uma proteína.

Regulação da expressão gênica eucariótica em nível traducional e pós-traducional

A regulação da expressão gênica em nível de *tradução* pode ocorrer pela ligação de fatores proteicos transatuantes a elementos reguladores cisatuantes presentes na molécula de mRNA, em geral presentes nas sequências não traduzidas 5′UTR e 3′UTR. Essa ligação pode facilitar ou perturbar a interação com outros fatores proteicos transatuantes, alterar a estrutura do RNA, possibilitar o encontro de sequências de RNA inicialmente distantes ou fornecer os sinais de localização para o transporte de moléculas de RNA. Além do mais, devemos ter em mente que para que a tradução aconteça é essencial que alguns elementos estejam disponíveis, tais como ribossomos, tRNA, fatores de iniciação e fatores de alongamento. Sendo assim, a disponibilidade desses elementos também está diretamente associada aos níveis de tradução. Por exemplo, nas respostas imunes a vírus é necessário a ativação de linfócitos T, que normalmente encontram-se no estágio G_0 do ciclo celular. A ativação dessas células se dá pelo aumento na síntese de proteínas que estimulam os linfócitos T a reingressarem no ciclo celular e proliferarem. O aumento na síntese de proteínas, por sua vez, é causado pelo aumento na disponibilidade de fatores de iniciação da tradução.

Outro exemplo de regulação da expressão gênica em nível de tradução ocorre durante o amadurecimento dos ovócitos humanos e nos estágios iniciais do desenvolvimento. Quando um ovócito humano é fertilizado, nenhum mRNA é inicialmente sintetizado até o estágio de 4

a 8 células, quando a transcrição do zigoto pode ser então iniciada. Nesse intervalo de tempo, as funções das células são determinadas pelos mRNA maternos sintetizados durante a ovocitogênese. Esses mRNA ficam armazenados nos ovócitos, sob uma forma inativa conferida por caudas poli (A) curtas. Na fertilização ou em uma fase posterior do desenvolvimento, os mRNA armazenados são ativados por um processo de poliadenilação citoplasmática semelhante à que ocorre no núcleo celular, tornando-os ativos.

A decodificação do mRNA durante a tradução resulta em proteínas novas que irão ocupar seu lugar no proteoma celular em substituição a proteínas que já chegaram ao final de sua vida útil. Para que isso aconteça, é necessário que essas proteínas remanescentes sejam devidamente removidas de modo *seletivo* para que não haja risco de degradação das proteínas recém-sintetizadas. Somado a isso, a degradação deve ser suficientemente rápida para possibilitar mudanças abruptas em resposta a determinadas exigências celulares. Enquanto em bactérias os tmRNA são responsáveis por etiquetar as proteínas sintetizadas de maneira incorreta para a degradação, em eucariotos, a maioria dos sistemas de degradação de proteínas envolve a participação da *ubiquitina*. Essa proteína intracelular se conjuga a proteínas-alvo e tem a capacidade de degradá-las por uma estrutura chamada de *proteossomo*.

Devemos lembrar ainda que após a tradução, algumas proteínas sofrem *processamento pós-traducional* pela clivagem proteolítica do produto primário, pela modificação de aminoácidos específicos ou pela excisão das inteínas, visando ao funcionamento, à atividade, à sinalização e à estabilidade das proteínas. Durante essa etapa, a regulação da expressão gênica também pode atuar.

Mecanismo de regulação da expressão gênica por RNA de interferência

Outra maneira de controlar a expressão gênica previamente à tradução é por meio da ação de moléculas de *RNA de interferência*, já descritas ao longo deste capítulo. Essas moléculas podem levar à inativação gênica por quatro caminhos. Em primeiro lugar, o complexo RISC que contém um RNA de interferência (*miRNA* ou *siRNA*) pode parear com moléculas de mRNA-alvo, clivando-as em regiões próximas à metade do RNA de interferência a ele ligado. Esse caminho que leva à degradação do mRNA após a clivagem é chamado de *slicer activity*. Um segundo meio pelo qual um RNA de interferência leva ao silenciamento gênico é pela *inibição da tradução* após o pareamento do RNA de interferência com o mRNA-alvo. Vários mRNA têm múltiplos sítios de ligação de miRNA que, quando ligados simultaneamente, inibem a tradução de maneira mais efetiva. Alguns RNA de interferência têm a habilidade ainda de silenciar a *transcrição*, alterando a conformação da cromatina. Esses siRNA interagem com proteínas para formar um *complexo de silenciamento transcricional de RNA* (RITS), semelhante ao complexo RISC. A partir daí, o RNA de interferência direciona o complexo RITS até uma molécula de DNA complementar e atrai enzimas responsáveis por metilar as caudas de histonas. Com isso, as histonas metiladas passam a interagir melhor com o DNA, levando a uma subsequente compactação da cromatina, que, por sua vez, impede o acesso da maquinaria transcricional. Finalmente, o quarto caminho pelo qual uma molécula de RNA de interferência regula a expressão gênica é pela *ativação da degradação do mRNA*-alvo a ele complementar em um processo *independente* da *slicer activity*. Nesse caminho, a degradação do mRNA requer a participação da enzima Dicer e do complexo RISC.

Amplificação gênica

Para alguns genes, uma das maneiras de se aumentar a atividade gênica se dá pela ampliação do número de cópias do gene, processo conhecido como *amplificação gênica*. Nessa estratégia, o número de moldes de DNA para a síntese de RNA é aumentado para possibilitar que mais produto gênico seja produzido. Contudo, como podemos aumentar o número de cópias de um gene para que ele se expresse mais? O exemplo mais clássico de amplificação gênica é o que envolve os genes de rRNA em ovócitos de anfíbios. Como vimos ao longo deste capítulo, os genes de rRNA são os componentes estruturais dos ribossomos e em eucariotos quatro tipos de rRNA podem ser encontrados: 5S, 5,8S, 18S e 28S. No genoma de *Xenopus laevis*, enquanto os genes que codificam o rRNA 5S estão distribuídos por todos os cromossomos em

grandes quantidades, os genes ordenados lado a lado que codificam os rRNA de 5,8S, 18S e 28S apresentam um número de cópias bem inferior, estando concentrados apenas nas regiões organizadoras de nucléolos. Nessas regiões, as unidades transcricionais encontram-se repetidas consecutivamente.

Durante a ovocitogênese em *Xenopus laevis*, a transcrição dos rRNA necessita ser reforçada porque o óvulo deve acumular um número sustancial de ribossomos para garantir o desenvolvimento precoce do embrião logo após a fertilização. Com isso, para garantir a demanda de rRNA 5,8S, 18S e 28S, os genes que os codificam são amplificados em ovócitos pela criação de cópias extra cromossômicas desses genes. Nessa estratégia, são formados *pequenos DNA circulares fechados covalentemente* que carregam os genes dos três tipos de rRNA e que têm a capacidade de se *replicar*, produzindo múltiplas cópias. A transcrição dessas moléculas de DNA circulares é suficiente para atender a grande demanda dos rRNA 5,8S, 18S e 28S no óvulo. Paralelamente, muitos dos genes que codificam o rRNA 5S ao longo do genoma de *Xenopus laevis* são ativados especificamente nos ovócitos, garantindo o quantitativo de ribossomos necessários no início do desenvolvimento embrionário.

Todos esses níveis de regulação da expressão gênica abordados ao longo desse capítulo ilustram a complexidade orquestrada inerente a estes eventos, principalmente quando consideramos que vários genes podem estar submetidos a diferentes formas de controle. Em humanos, estes mecanismos têm uma importância particular ainda maior pelo fato de várias doenças serem atribuídas a um controle ineficiente da expressão gênica.

RESUMO

- Um gene pode ser definido como uma unidade funcional que controla a síntese de um produto final, que pode ser uma molécula de RNA ou um polipeptídio. Dois tipos de genes funcionais estão presentes no genoma humano: os que codificam moléculas de RNA e os que codificam proteínas
- Os genes que expressam RNA passam apenas pelo primeiro estágio da expressão gênica, a transcrição, enquanto a expressão dos genes que codificam proteínas envolve dois passos subsequentes (a transcrição e a tradução)
- Na transcrição, um segmento de nucleotídios de uma das fitas do DNA serve de molde (3′→5′) para a síntese de uma molécula de RNA complementar antiparalela (5′→3′) pela ação de uma RNA polimerase
- Para que um gene seja transcrito, é necessário que ele tenha elementos conservados em sua estrutura. Uma unidade de transcrição consiste em um promotor, a região codificante de RNA e um finalizador
- A transcrição de procariotos difere da de eucariotos em alguns aspectos. Enquanto as bactérias têm apenas uma RNA polimerase, eucariotos têm 3 tipos da enzima (RNA polimerases I, II e III) com especificidades distintas. Em bactérias, as regiões promotoras são reconhecidas diretamente pela RNA polimerase, enquanto em eucariotos essas enzimas não têm autonomia para iniciar a transcrição, necessitando do auxílio de fatores proteicos adicionais. Os promotores de eucariotos são mais complexos e consistem não somente da sequência na qual a RNA polimerase se liga ao DNA, como de outras sequências que atuam como reguladoras da expressão gênica, inibindo ou estimulando a transcrição. Somado a isso, em eucariotos, o início da transcrição requer a modificação da estrutura da cromatina
- Vários tipos de RNA podem ser produzidos pela transcrição, dos quais os principais são o mRNA, o rRNA e o tRNA. Após a síntese, esses RNA devem ser processados antes de executarem suas funções celulares
- O processamento do pré-mRNA eucariótico envolve a introdução de um nucleotídio modificado na extremidade 5′ do transcrito (capeamento 5′), a poliadenilação da extremidade 3′ e a excisão dos íntrons seguida do encadeamento dos éxons. Tanto o capeamento quanto a poliadenilação são essenciais para a estabilidade do mRNA maduro no transporte do núcleo para o citoplasma e para o início da tradução. O mRNA bacteriano é traduzido imediatamente após a tradução e sofre pouco processamento
- Na tradução, uma molécula de mRNA é decodificada para especificar a síntese de um polipeptídio formado pela união de vários aminoácidos ligados em série. A sequência de aminoácidos é determinada pela sequência de nucleotídios presente no mRNA, enquanto as moléculas de tRNA são as responsáveis por reconhecer a sequência do mRNA e correlacioná-la com os aminoácidos correspondentes. O processo acontece nos ribossomos, que posicionam corretamente os tRNA com o mRNA, catalisando as ligações entre os aminoácidos que são adicionados ao polipeptídio nascente. Os ribossomos são formados por várias moléculas de rRNA e várias proteínas diferentes
- As regras que regem a leitura da molécula de mRNA e a decodificação dos nucleotídios em aminoácidos durante a tradução compõem o código genético. Esse código é lido em trincas (códons), é degenerado e quase que completamente universal
- As células têm mecanismos de vigilância do mRNA que possibilitam eliminar erros que podem interferir na tradução
- As bactérias utilizam o princípio da economia máxima de energia e da flexibilidade bioquímica para regular seus genes,

de modo que somente serão expressos aqueles produtos gênicos necessários para cada condição presente no ambiente. Nesses organismos, os genes estruturais são agrupados em óperons, sendo transcritos juntos como uma única molécula de RNA

- Nos óperons indutíveis, a transcrição está inativa e uma mudança deve acontecer para estimular a transcrição. Nos óperons repressíveis, a transcrição está ativa e uma mudança deve acontecer para inibir a transcrição. No controle negativo, o produto do gene regulador é uma proteína reguladora repressora que se liga ao operador e inibe a transcrição dos genes estruturais. No controle positivo, o produto do gene regulador é uma proteína reguladora ativadora, que é necessária para induzir a transcrição dos genes estruturais
- Outras formas de regulação gênica em procariotos envolvem a atenuação, o controle sobre a tradução e a regulação pós-traducional pela inibição por *feedback*
- Em eucariotos, a complexidade da regulação gênica é aumentada pela necessidade de especialização dos diferentes tecidos, pela expressão temporalmente diferenciada dos genes de eucariotos e pela necessidade de resposta a estímulos externos
- A regulação da expressão gênica em eucariotos pode acontecer em diferentes níveis que incluem a alteração dinâmica da estrutura da cromatina, transcrição, processamento do mRNA, estabilidade do mRNA, tradução e pós-tradução. Grande parte do controle gênico se dá pela atuação de proteínas reguladoras que se ligam a regiões específicas do DNA.

AUTOAVALIAÇÃO

2.1 Comente quais estruturas um gene precisa ter para ser funcional.

2.2 Por que a definição clássica de um gene, na qual ele corresponderia a uma unidade funcional capaz de gerar uma molécula de RNA ou um polipeptídio, pode não ser real para todas as situações?

2.3 Descreva as diferentes etapas que compõem o processo transcricional em bactérias.

2.4 Quais características diferem a transcrição em eucariotos da que ocorre em procariotos?

2.5 Distinga fatores transatuantes de elementos cisatuantes.

2.6 Como podemos classificar os promotores utilizados na transcrição de genes codificadores de proteínas em eucariotos?

2.7 Quais os principais tipos de RNA produzidos durante a transcrição e quais os seus papéis funcionais?

2.8 Descreva o processamento do pré-mRNA e qual a importância das diferentes etapas que compõem este processamento.

2.9 Em que consiste o encadeamento alternativo dos éxons durante o processamento do pré-mRNA?

2.10 Como os pequenos RNA de interferência são processados?

2.11 Descreva as principais características do código genético.

2.12 Em que consistem as três etapas que fazem parte do processo de tradução?

2.13 Como podemos diferenciar a tradução de eucariotos da que ocorre em procariotos?

2.14 Em que consiste o processo de vigilância do mRNA e como ele acontece em bactérias?

2.15 De que maneira uma mutação sinônima pode ter efeito sobre a tradução, mesmo que não afete os sítios reguladores do processo de recomposição dos éxons?

2.16 O que são modificações pós-traducionais?

2.17 Para que serve a regulação da expressão gênica?

2.18 Distinga óperons indutíveis de repressíveis e controle negativo de positivo.

2.19 Comente como funciona o óperon Lac.

2.20 Defina o que é um mecanismo de regulação gênica por atenuação.

2.21 Como um RNA antissenso pode regular a expressão de um gene?

2.22 Defina o conceito de fatores epigenéticos e como eles podem afetar a regulação da expressão gênica.

2.23 Por que podemos dizer que as histonas têm um código genético próprio?

2.24 Como acontece a inativação de um dos cromossomos X nas mulheres e qual o seu objetivo?

2.25 Em que consiste o *imprinting* parental?

2.26 Defina controle da expressão gênica de longo alcance.

2.27 O que são proteínas de choque térmico?

2.28 Quais mecanimos podem acontecer durante o processamento do pré-mRNA de eucariotos capazes de regular a expressão gênica?

2.29 De que maneiras um RNA de interferência pode reprimir a expressão de um gene?

2.30 Defina amplificação gênica, citando um exemplo de sua atuação.

BIBLIOGRAFIA

ALBERTS, B.; JOHNSON, A.; LEWIS, J. *et al*. *Biologia Molecular da Célula*. 5ª ed., Porto Alegre: Artmed, 2010.

BERLETCH, J.B.; YANG, F.; XU, J. *et al*. Genes that escape from X inactivation. *Hum Genet*. 2011; 130:237-45.

CAMPBELL, N.A.; REECE, J.B.; URRY, L.A. *et al*. *Biologia*. 8ª ed., Porto Alegre: Artmed, 2010.

CHAMARY, J.V.; HURST, L.D. The price of silent mutations. *Scientific American* 2009; 46-53.

CHOW, J.C.; BROWN, C.J. Forming facultative heterochromatin: silencing of an X chromosome in mammalian females. *Cell Mol Life Sci*. 2003; 60:2586-603.

DEANE, C.M.; SAUNDERS, R. The imprint of codons on protein structure. *Biotechnol J*. 2011;6:641-9.

GONZALEZ, S.; PISANO, D.G.; SERRANO, M. Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. *Cell Cycle* 2008; 7:2601-8.

IBORRA, F.J.; JACKSON, D.A.; COOK, P.R. The case for nuclear translation. *J Cell Sci*. 2004;117:5713-20.

KALANTRY, S. Recent advances in X-chromosome inactivation. *J Cell Physiol*. 2011;226:1714-8.

LEWIN, B. *Genes IX*. 9ª ed., Porto Alegre: Artmed, 2009.

NIELSEN, H.; JOHANSEN, S.D. Group I introns: moving in new directions. *RNA Biol* 2009; 6: 375-83.

PIERCE, B.A. *Genética: um enfoque conceitual*. 3ª ed., Rio de Janeiro: Guanabara Koogan, 2011.

ROSS, M.T.; GRAFHAM, DV.; COFFEY, A.J. *et al*. The DNA sequence of the human X chromosome. *Nature* 2005; 434: 325-337.

SANTOS-REBOUÇAS, C.B.; PIMENTEL, M.M. Implication of abnormal epigenetic patterns for human diseases. *Eur J Hum Genet*. 2007;15:10-7.

SANTOS-REBOUÇAS, C.B.; PIMENTEL, M.M. MicroRNAs: macro challenges on understanding human biological functions and neurological diseases. *Curr Mol Med*. 2010;10:692-704.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética*. 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

STRACHAN, T.; READ, A.P. *Genética molecular humana*. 2ª ed., Porto Alegre: Artmed, 2002.

TOOR, N.; KEATING, K.S.; PYLE, A.M. Structural insights into RNA splicing. *Curr Opin Struct Biol*. 2009;19:260-6.

TURAN, N.; KATARI, S.; GERSON, L.F. *et al*. Inter and intraindividual variation in allele-specific DNA methylation and gene expression in children conceived using assisted reproductive technology. *PLoS Genet*. 2010; 6:e1001033.

3

Mutações no DNA e Mecanismos de Reparo

Objetivos de estudo

Saber reconhecer os principais tipos de mutações gênicas, bem como identificar os mecanismos capazes de originá-las

Entender as consequências das diferentes classes de mutações para o funcionamento gênico e para o organismo como um todo

Compreender como mutações em regiões não codificantes podem resultar em um fenótipo anormal

Entender os conceitos de mutação dinâmica e de antecipação genética

Ser capaz de reconhecer as peculiaridades das mutações no genoma mitocondrial

Saber utilizar a nomenclatura internacional para a descrição de alterações gênicas

Entender como os diferentes agentes mutagênicos químicos e físicos podem induzir o aparecimento de mutações

Compreender os principais mecanismos de reparo descritos e como alterações nesses mecanismos podem levar a doenças

Conceitos-chave do capítulo

Agente mutagênico
Agentes alquilantes
Agentes intercalantes
Análogos de base
Antecipação genética
Códon
Códon genético
Decaimento de mRNA
Deleção gênica
Duplicação gênica
Farmacogenômica
Heteroplasmia
Homoplasmia
Hot spot
Inserção gênica
Mecanismo de reparo
Microssatélite
Mutação
Mutação de sentido trocado
Mutação dinâmica
Mutação espontânea
Mutação *frameshift*
Mutação germinativa
Mutação *missense*
Mutação mitocondrial
Mutação neutra
Mutação *nonsense*
Mutação nula
Mutação pontual
Mutação reversa
Mutação sem sentido
Mutação *sense*
Mutação silenciosa
Mutação somática
Mutação supressora
Nomenclatura
Pareamento incorreto por deslize das fitas
Polimorfismo
Reparo de mau pareamento
Reparo por excisão de base
Reparo por excisão de nucleotídios
Reparo por recombinação
Sítios de encadeamento
Transição
Transversão

Introdução

Mutação
Alteração na sequência de nucleotídios de um gene

Neste capítulo, abordaremos os principais tipos de mutação conhecidos, suas causas e consequências, bem como os mecanismos existentes para reparar cada um desses tipos. O **termo** mutação pode ser definido como qualquer alteração ocorrida na molécula de DNA durante sua replicação e que não tenha sido devidamente corrigida pelos sistemas de reparo existentes em nosso organismo.

Apesar de ser um evento complexo, em geral, a replicação do DNA é um mecanismo que mantém fidelidade muito alta à sequência molde da molécula de DNA. A DNA polimerase, **principal enzima envolvida no processo, se desloca a uma velocidade de 20 pares de base por** segundo durante a replicação do DNA e introduz um nucleotídio errado a cada 10 milhões de pares de bases. Isso significa dizer que, ao longo da leitura deste capítulo, você sofrerá milhares de mutações em suas células. Contudo, isso não é motivo para preocupação, visto que mais de 99,9% dos erros ocorridos durante a replicação do DNA são corrigidos por sistemas enzimá-

ticos específicos, os quais serão abordados adiante. Além disso, não podemos esquecer que a incorporação de variações ao longo de nosso genoma constitui grande fonte de variabilidade genética, componente essencial ao processo evolutivo. Assim, a baixa frequência na introdução de uma mutação pode ser encarada como um equilíbrio entre o surgimento de uma novidade evolucionária em detrimento da ocorrência de uma doença ou morte em uma porcentagem de membros de nossa espécie.

Embora uma mutação possa ocorrer em qualquer sítio ao longo de nosso genoma, existem regiões mais propensas a erros durante a replicação, chamadas de pontos quentes ou *hot spots*.

Polimorfismo
Existência de uma variante na sequência de nucleotídios de um gene que apresente frequência populacional igual ou superior a 1%

Quando uma variante surge como uma mutação e se torna comum, fixando-se na população com uma frequência de, no mínimo, 1%, ela recebe o nome de polimorfismo. Com isso, um único loco passa a ter mais de um alelo funcional. Essas variantes levam a características próprias de cada indivíduo e nos diferem parcialmente, tendo hoje aplicações muito relevantes na confecção de medicamentos personalizados pelo estudo da Farmacogenômica. Em geral, os polimorfismos não são patogênicos, mas podem modular o risco de desenvolvimento de determinadas doenças humanas. Além da frequência populacional, outros fatores que distinguem um polimorfismo de uma variante patogênica incluem a análise da conservação evolutiva e a presença entre membros de uma mesma família. Caso a sequência de DNA modificada seja conservada entre espécies, ou seja, tenha alta ortologia, isso significa que a variante introduzida pode estar associada a um fenótipo anormal. Além disso, uma mudança de sequência presente em um indivíduo afetado por dada doença, mas não em seus pais não afetados, pode ser patogênica.

Causas, tipos e efeitos das mutações

Quanto à sua natureza, podemos classificar as mutações em dois subtipos: *espontâneas* e *induzidas*. As mutações espontâneas surgem *ao acaso*, em decorrência de erros normalmente introduzidos durante o processo de replicação do DNA, sem que haja interferência de qualquer agente capaz de provocá-las.

A *taxa de mutação espontânea* da espécie humana, estimada por sequenciamento direto, gira em torno de 10^{-8} mutações por loco por geração, mas alguns genes podem apresentar maior frequência de erros durante a replicação.

Um exemplo clássico é o gene da distrofina humana, que está localizado no cromossomo X e tem uma taxa de mutação de 10^{-4} mutações por loco por geração. Acredita-se que a alta taxa de mutações associada a esse gene se deva a sua grande extensão, que encobre 2,4 milhões de pares de bases (em média, os genes humanos têm entre 10 e 15 mil pares de bases), o que lhe garante o título de maior gene humano conhecido até o momento, levando 16 h para ser transcrito.

Diferentemente das mutações espontâneas, as mutações induzidas ocorrem pela ação de agentes físicos e/ou químicos, denominados *agentes mutagênicos*, que podem ser encontrados no ambiente externo ou podem ser até mesmo gerados no ambiente intracelular.

Os principais *agentes mutagênicos* são: físicos (luz ultravioleta e radiações ionizantes) e químicos (análogos de bases, agentes intercalantes e produtos reativos internos).

A grande fonte de mutações decorre de erros espontâneos ocorridos na replicação e no reparo da molécula de DNA. Em hominídios, a frequência de mutações deletérias carregadas por cada um de nós a cada geração varia de 2,5 a 3 mutações. Você deve se perguntar: e por que não estamos extintos diante da presença de tais mutações? Simplesmente porque a reprodução sexuada possibilita que a recombinação meiótica de nossos gametas ocorra a cada geração e elimine as mutações mais prejudiciais.

Mutações somáticas
Alterações na sequência do DNA em células do corpo que não incluem os gametas

Espontânea ou induzida, a mutação pode ocorrer em qualquer célula de nosso organismo, tanto nas células somáticas, dispostas na maioria de nossos tecidos, quanto nas *células germinativas*, que originarão nossos gametas.

Mutações germinativas
Alterações na sequência do DNA em células que originarão os gametas

As mutações somáticas causam danos apenas à própria pessoa, e não introduzem novas formas alélicas na população. Contrariamente, nas mutações germinativas, caso a alteração não prejudique a capacidade reprodutiva do indivíduo, poderá ser transmitida para as próximas gerações,

causando doenças hereditárias, principalmente no caso de mutações germinativas dominantes. O fato de as mutações somáticas não se propagarem ao longo das gerações não subestima sua relevância, visto que esse tipo de mutação é responsável por um vasto número de tipos de câncer e pelas inúmeras alterações que podem surgir no intenso mecanismo celular que acontece desde a formação do zigoto até o desenvolvimento completo do embrião.

Devemos notar ainda que as frequências de mutações que acontecem nas linhagens germinativas são distintas para homens e mulheres em virtude das peculiaridades do processo de gametogênese. Os gametas femininos já estão todos pré-formados ao nascimento e ficam estacionados em meiose I (Capítulo 4). Posteriormente, somente duas divisões celulares ocorrem da puberdade até a menopausa para produzirem o ovócito maduro, sendo que o número total estimado de divisões celulares femininas das células germinativas até o óvulo maduro gira em torno de 24 a 31. Muitos ovócitos ficam estacionados durante décadas, o que aumenta o risco de não disjunções cromossômicas, como as trissomias dos cromossomos 13, 18 e 21, nos embriões de mulheres que engravidam com mais de 35 anos. Por outro lado, ao se tratar de mutações gênicas, aparentemente essas alterações são preferencialmente herdadas pela via patrilinear. Os gametas masculinos são constantemente renovados após a puberdade e são produzidos em quantidades muito maiores do que os femininos. Ocorrem cerca de 30 divisões celulares (precedidas por replicações) do desenvolvimento intrauterino até a puberdade e, a partir deste período, em torno de 23 ciclos ocorrem por ano para a produção dos espermatozoides. Considerando-se a idade média reprodutiva de 25 anos para os homens, seus gametas já teriam sofrido um número de divisões celulares muito superior ao dos gametas femininos (por volta de 310 divisões celulares), sendo mais suscetíveis a mutações gênicas.

A maioria das mutações que acontecem em nosso organismo não é fixada na população.

Além da fidedignidade de nossa maquinaria replicacional e da atuação eficiente dos sistemas de reparo, outra maneira de filtrar as mutações introduzidas em nosso genoma é pela inviabilidade celular. Isso significa que, caso a mutação cause letalidade para aquela célula que a apresenta, esta não será perpetuada entre as células-filha, não ocasionando consequências deletérias em nível somático e/ou gamético. Outra questão interessante é que mesmo que determinado gene mutado tenha expressão ubíqua, em geral apenas um subconjunto dos tecidos que o expressam é afetado. Isso pode ocorrer por dois motivos principais. Primeiramente, devemos notar que os vários tipos celulares apresentam papéis e necessidades metabólicas distintos, podendo ser diferencialmente vulneráveis a uma mesma mutação. Além disso, diferentes níveis de redundância funcional estão presentes em cada tipo celular, ou seja, em determinado tecido, a expressão de outro(s) gene(s) com funcionalidade semelhante à daquele que está mutado pode suprir sua disfunção, ao passo que em outros tecidos não. Quando uma mutação elimina completamente a função de um gene, ela é dita mutação nula. Mutações nulas ou alterações que impedem o gene de exercer a sua função, mesmo que parcialmente, são nomeadas *mutações de perda de função*, as quais em geral são recessivas. No entanto, quando uma mutação faz com que a proteína produzida tenha uma nova função ou uma função exacerbada, a mudança é chamada de *mutação de ganho de função* e, em geral, tem padrão dominante.

Mutação nula
Alteração na sequência de um gene que abole completamente a sua função na célula

Há quatro principais tipos de mutações gênicas: *mutações de ponto*, *deleções*, *inserções* e *inversões*, e o modo como a expressão do gene mutante é alterada depende do tipo de mutação e de sua localização ao longo de dado gene.

Como ilustrado na Figura 3.1, as *mutações de ponto* consistem na substituição de um nucleotídio por outro na molécula de DNA. Em casos mais raros, uma sequência contígua de bases pode ser substituída simultaneamente como resultado de um mecanismo chamado de *conversão gênica*, comum ao cromossomo Y (Capítulo 4). Quando a mutação de ponto envolver a troca de uma purina (adenina e guanina) por outra purina ou de uma pirimidina (citosina e timina) por outra pirimidina, a mutação de ponto é chamada de *transição*. Quando a substituição for realizada entre uma purina e uma pirimidina ou vice-versa, a mutação de ponto recebe o nome de *transversão*. Como a transversão envolve 8 possibilidades de substituição (A↔C, A↔T, C↔G, T↔G) e a transição, apenas quatro eventos possíveis de troca (A↔G e C↔T), poderíamos esperar que a frequência de transversões fosse maior no genoma em comparação à frequência de transições. Contudo, o alto número de transições de citosina para timina em

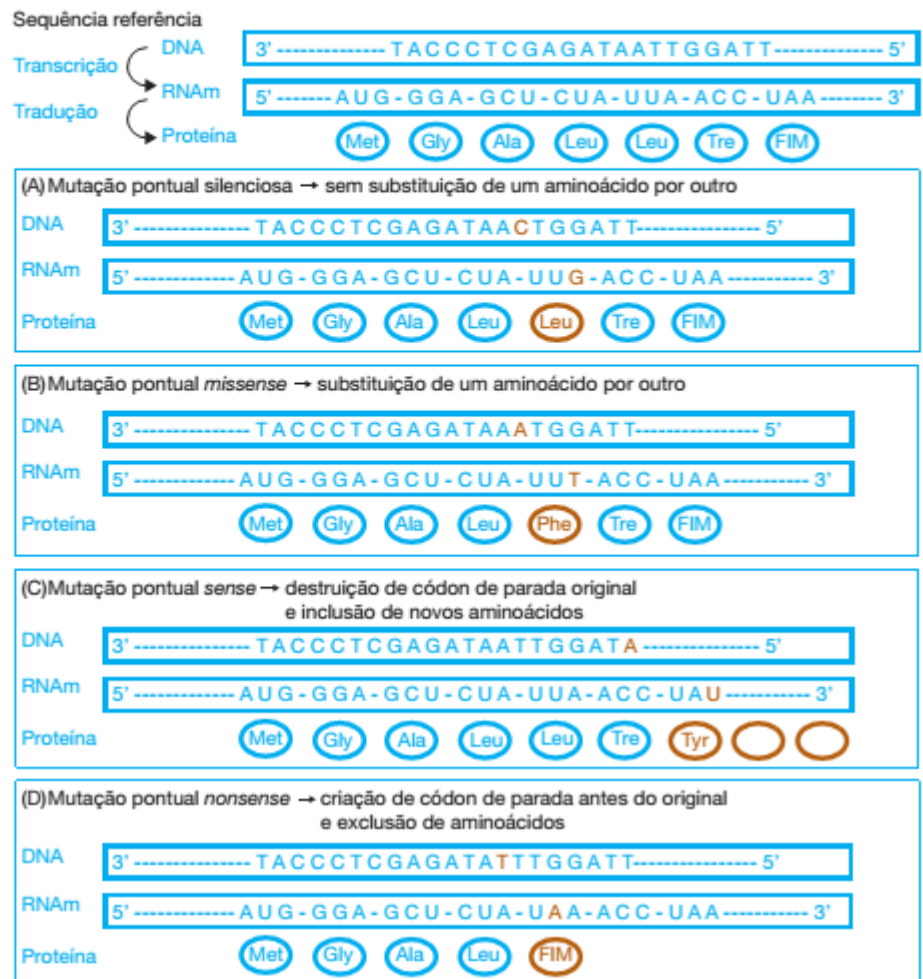



Figura 3.1 Mutação pontual e seus possíveis efeitos. (A) Mutação pontual silenciosa; (B) Mutação *missense*; (C) mutação *sense*, representada pela substituição pontual com destruição do códon de parada original; (D) mutação *nonsense* levando a códon de parada prematuro.

nosso genoma faz com que, na realidade, o observado seja exatamente o oposto. Isso se deve às inúmeras citosinas presentes nos dinucleotídios CG ligados covalentemente (ilhas CpG), os quais são altamente suscetíveis à adição de um grupamento metil a seus carbonos 5′ pela ação de DNA metil transferases. Após o processo, as citosinas metiladas se tornam quimicamente instáveis e tendem a sofrer desaminação espontânea, sendo transformadas estruturalmente em timinas. Assim, as citosinas metiladas presentes nos dinucleotídios CG são consideradas um bom exemplo de *hot spots* no nosso genoma.

A partir do projeto genoma humano, foi constatado que as substituições de nucleotídio são muito mais comuns ao longo de nosso genoma do que se imaginava, afetando um nucleotídio a cada 200 pb a 500 pb. Essas substituições de nucleotídio único passaram a ser chamadas de SNP (do inglês, *single nucleotide polymorphism*), sendo consideradas responsáveis em parte pelas diferenças fenotípicas interindividuais. Atualmente, mais de 187 milhões de SNP presentes no genoma humano foram submetidos ao banco de dados dbSNP (http://www.ncbi.nlm.nih.gov/snp) em parte como resultados obtidos pelos projetos 1000 Genomes e ClinSeq Sequencing Project. A identificação de SNP que estejam associados a uma resposta diferenciada a medicamentos e à predisposição a doenças pode trazer novos rumos para a medicina individualizada. Além disso, a identificação dessas variações por meio de ferramentas de investigação molecular em larga escala pode fornecer um panorama sobre nossa história evolutiva.

Código genético
Conjunto de regras que governa a tradução de uma molécula de mRNA em uma proteína

Códon
Sequência de três nucleotídios adjacentes que especificam um aminoácido

Pelo fato de nosso código genético ser degenerado (Capítulo 2), ou seja, mais de um códon dar origem a um mesmo aminoácido, uma mutação de ponto pode levar ou não à substituição de um aminoácido por outro na cadeia polipeptídica. Quando o aminoácido não se altera,

chamamos a mutação de ponto de *silenciosa*, ao passo que, quando a substituição altera o aminoácido que seria codificado pelo códon original, a mutação se chama ***missense*** **ou** mutação de sentido trocado. Nas mutações ***missense***, os efeitos deletérios dependerão das distinções químicas e funcionais entre o aminoácido original e o substituinte, já que algumas substituições podem provocar troca de carga elétrica e de polaridade, causando modificações expressivas na estrutura tridimensional da proteína e em sua função. Entretanto, não se engane em achar que mutações silenciosas são sempre inofensivas. Lembre-se de que mesmo que não tenha havido troca de aminoácidos, se a substituição de nucleotídios ocorrer em sequências evolutivamente conservadas envolvidas em mecanismos regulatórios, como em sítios exônicos participantes como acentuadores do *splicing* (*exonic splicing enhancers* ou ESE), os efeitos poderão ser extremamente deletérios. Até o momento, cerca de 50 doenças, incluindo a fibrose cística, foram associadas em parte a mutações conhecidas como silenciosas, das quais várias atuam prejudicando a remoção dos íntrons. No entanto, a perturbação do *splicing* não é a única maneira de uma mutação silenciosa levar a doenças. Mesmo que os íntrons sejam removidos adequadamente, a presença de uma alteração silenciosa na sequência do mRNA, pode interferir no pareamento intramolecular de sequências complementares na molécula de RNA. A dobradura do mRNA é um dos mecanismos que determina sua estabilidade. Além disso, sabe-se que embora o código genético seja degenerado, códons sinônimos podem não ser utilizados em iguais proporções pela maquinaria traducional e essa escolha preferencial pode influenciar a taxa e a acurácia da síntese proteica. Nesse sentido, um códon que pareie com tRNA mais abundantes pode ser traduzido mais rapidamente (Capítulo 2).

Mutação de sentido trocado
Mutação que altera o aminoácido original codificado anteriormente pelo códon

Mutação silenciosa
Aquela que não altera o aminoácido original codificado por um dado códon

Curiosamente, dentre as três bases que compõem um códon, a frequência de substituições na segunda é sempre a mais baixa, provavelmente em decorrência de uma pressão seletiva conservativa, visto que mutações nesse sítio sempre causam a troca de aminoácido. Quando se altera a sequência de aminoácidos da proteína, mas não a sua função, há a mutação neutra.

Mutação neutra
Mutação que altera a sequência de aminoácidos de uma proteína sem ter efeito sobre sua funcionalidade

A mutação de ponto também pode alterar o códon de iniciação (AUG) ou criar e destruir um dos três códons de parada utilizados durante a tradução do mRNA (UAA, UAG, UGA).

Quando o códon que originalmente especificaria um aminoácido é convertido em um códon de parada, há a mutação sem sentido **ou** ***nonsense***. O produto truncado ocasionado pode ser funcionalmente anormal, interferindo bruscamente nas funções celulares, ou ser instável a ponto de levar a sua consequente degradação, muitas vezes ainda em nível de mRNA. Nesse sentido, a **mutação *nonsense*** pode ser alvo de um mecanismo de sobrevivência chamado de ***decaimento de mRNA mediado por mutações nonsense*** (do inglês, ***nonsense-mediated decay***). Por meio desse mecanismo, o mRNA mutado é degradado antes de ser traduzido, ocasionando perda total de função da proteína em detrimento de uma função anormal desta na célula. É estimado que cerca de um terço de todas as doenças genéticas estejam relacionadas a esse mecanismo.

Mutação sem sentido
Mutação que transforma uma sequência codificante em um códon de término da tradução

Da mesma maneira que a mutação pontual pode levar a um códon prematuro de término da tradução, um dos códons finalizadores pode ser convertido por uma mutação pontual em um códon que especifica um aminoácido, dando origem a uma mutação *sense*. Nesse caso, o ribossomo pode chegar o final do mRNA sem encontrar o códon de término, e devido a falta de sinalização para finalizar a tradução, ele permanece ligado ao mRNA. A permanência do ribossomo no mRNA impede que este ribossomo seja reutilizado em uma nova tradução e, caso o evento ocorra de modo recorrente, pode afetar os níveis globais de síntese de proteínas.

Mutação *sense*
Mutação que transforma um códon finalizador em um códon que especifica um aminoácido

Um exemplo de mutação de ponto que ocasione efeito fenotípico aparente é a anemia falciforme, que ocorre por causa da substituição de uma adenina por uma timina no gene da β-globina localizado no cromossomo 11, levando à troca de um ácido glutâmico por uma valina na extremidade amino da proteína. Com essa modificação, a hemoglobina mutada passa a se cristalizar nas hemácias, fazendo com que estas adquiram um formato de foice (falciforme) em condições de baixa tensão de oxigênio. Essas células alteradas obstruem os pequenos vasos sanguíneos, impedindo a chegada de oxigênio nos diferentes tecidos. Em regiões endêmicas para a malária, como ocorre em vários países africanos, os heterozigotos portadores de tal mutação apresentam vantagem seletiva sobre o seu agente causador (*Plasmodium falciparum*) em relação aos homozigotos normais, que morrem da malária, e os indivíduos homozigotos falcêmicos, que costumam não sobreviver à grave anemia na infância. Esse fato demonstra a interferência adaptativa constante do meio ambiente sobre a fixação de uma nova mutação.

Uma variante que é bem-sucedida em um meio, não necessariamente o será em um ambiente diferente.

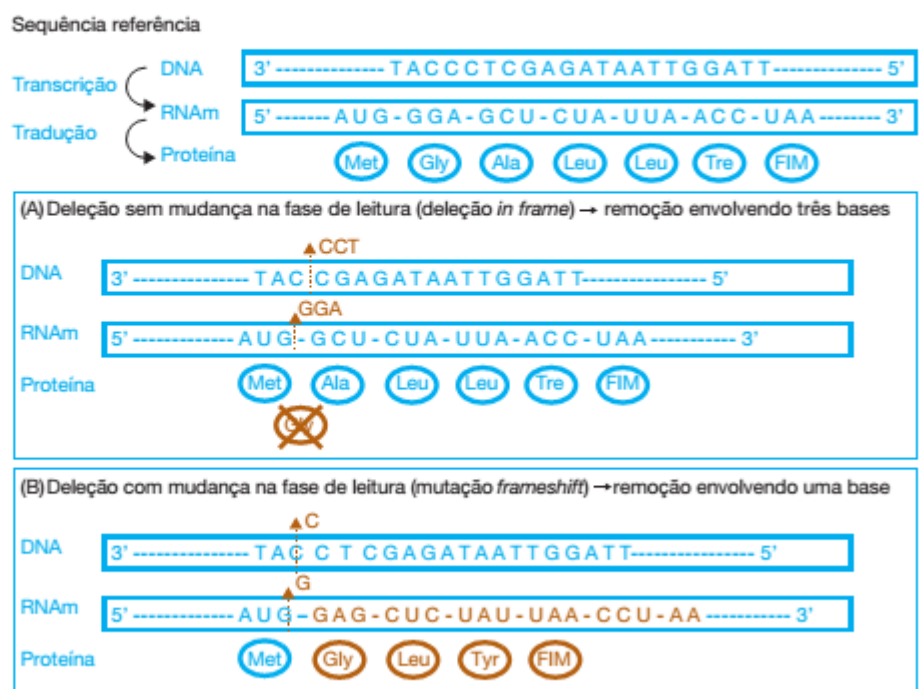
As *deleções* são tipos de mutações que consistem na excisão de um ou mais pares de bases na sequência do DNA (Figura 3.2). Como abordado no Capítulo 2, os nucleotídios que compõem o mRNA maduro são lidos, durante a tradução, em unidades de três nucleotídios, chamadas de códons.

Quando o número de bases em uma deleção não é múltiplo de três, a mutação altera a matriz ou fase de leitura dos códons durante a tradução do RNAm, resultando em uma sequência de aminoácidos distinta da original, desde o ponto da deleção (mutação *frameshift*).

Mutação *frameshift*
Alteração na sequência de nucleotídios de um gene que leva à mudança na matriz de leitura dos seus códons

Com isso, certamente podemos dizer que tamanho não é documento, já que a deleção de uma única base pode ser muito mais prejudicial que outra que envolva, por exemplo, 9 bases consecutivas. Assim como acontece nas mutações pontuais, uma deleção *frameshift* pode ainda alterar códons de iniciação ou de parada. Dessa maneira, os efeitos nocivos de uma deleção dependerão da extensão da mutação, se houve alteração da matriz de leitura de determinado gene e da importância dos aminoácidos envolvidos na funcionalidade proteica.

Para exemplificarmos os efeitos que uma deleção pode causar, voltemos ao gene da distrofina humana, importante proteína presente na musculatura estriada esquelética que atua em nível de membrana e funciona como conectora do citoesqueleto, garantindo a contração muscular. A ausência dessa proteína leva à perda progressiva da capacidade de contração muscular. Assim, mutações nesse gene estão associadas a duas formas distintas de distrofia muscular. Em sua forma mais branda (distrofia muscular de Becker), em geral as deleções não alteram a matriz de leitura dos códons do gene, possibilitando a produção de uma proteína que, embora seja menor em tamanho que a original, ainda mantém um pouco de sua funcionalidade e provoca um quadro de degeneração muscular mais tênue. Quando a deleção é do tipo *frameshift*, uma proteína totalmente distinta passa a ser produzida no lugar da distrofina e isso resulta em um quadro clinicamente mais grave da doença (distrofia muscular de Duchenne), o qual ocasiona morte em seus portadores ainda na adolescência ou no início da fase adulta.

Ao contrário das deleções, as *inserções* são mutações gênicas que correspondem à inclusão de uma ou mais bases na sequência do DNA durante a sua replicação (Figura 3.3). Do mesmo modo, quando o número de bases adicionadas não é múltiplo de 3, a inserção altera a matriz de leitura dos códons durante a tradução, o que resulta em uma sequência de aminoácidos

Sequência referência
Transcrição — DNA: 3' -------------- T A C C C T C G A G A T A A T T G G A T T -------------- 5'
RNAm: 5' ------- A U G - G G A - G C U - C U A - U U A - A C C - U A A -------- 3'
Tradução — Proteína: Met Gly Ala Leu Leu Tre FIM

(A) Deleção sem mudança na fase de leitura (deleção *in frame*) → remoção envolvendo três bases
CCT
DNA: 3' --------------- T A C C G A G A T A A T T G G A T T---------------- 5'
GGA
RNAm: 5' ------------- A U G - G C U - C U A - U U A - A C C - U A A ----------- 3'
Proteína: Met Ala Leu Leu Tre FIM
Gly

(B) Deleção com mudança na fase de leitura (mutação *frameshift*) → remoção envolvendo uma base
C
DNA: 3' --------------- T A C C T C G A G A T A A T T G G A T T---------------- 5'
G
RNAm: 5' ------------- A U G – G A G - C U C - U A U - U A A - C C U - A A ----------- 3'
Proteína: Met Gly Leu Tyr FIM

Figura 3.2 **Tipos de deleção gênica e seus possíveis efeitos.** (A) Deleção de um número de bases múltiplo de 3, sem mudança da matriz ou fase de leitura; (B) mutação *frameshift*, na qual a deleção de uma base provoca mudança na fase de leitura e, nesse caso, ainda provocou a criação de um códon de parada prematuro.

Figura 3.3 Inserção gênica e seus possíveis efeitos. (A) Inserção de um número de bases múltiplo de 3 sem mudança da matriz ou fase de leitura; (B) inserção de apenas uma base, provocando a mudança na fase de leitura e destruição do códon de parada, fazendo com que a proteína seja traduzida até que um novo códon de término seja alcançado.

totalmente distinta a partir do ponto de mutação. Um códon de iniciação ou parada também pode ser alterado. Regiões de alta similaridade entre suas sequências de nucleotídios funcionam como sítios de recombinação homóloga desigual, determinando o surgimento de inserções e deleções. Inserções gênicas também podem ocorrer por meio de um intermediário de RNA.

A doença de Tay-Sachs, uma disfunção autossômica recessiva muito comum na população judaica Asquenaze, resulta de uma deficiência na síntese da enzima hexosaminidase A. Na maioria dos casos, uma inserção de 4 bases no éxon 11 do gene *HEXA*, situado no cromossomo 15, leva a um códon finalizador prematuro. Na ausência da enzima hexosaminidase A, a GM2 gangliosidase não é hidrolisada e se acumula principalmente no tecido neuronal, resultando em degeneração neuronal progressiva. Em geral, crianças com a doença começam a manifestar os primeiros sintomas a partir dos 6 meses, quando passam a exibir atraso progressivo no desenvolvimento, dificuldades respiratórias, eventual cegueira e deficiência intelectual, culminando em óbito geralmente antes dos 5 anos.

É válido ressaltar que deleções ou inserções que não afetam a matriz de leitura gênica são ditas "*in frame*". Contudo, mesmo que mantenham a matriz de leitura intacta, a remoção ou adição de um ou mais aminoácidos pode afetar o fenótipo.

O próximo tipo de alteração que vamos abordar é a *inversão gênica*, que consiste na excisão de um segmento da sequência do DNA, seguida de sua reinserção na mesma posição, mas em orientação inversa (Figura 3.4). Embora a inversão gênica não ocasione mudanças na fase de leitura de um gene, ela pode destruir o códon de iniciação e criar ou destruir códons de parada, tendo consequências fenotípicas graves. A hemofilia A, causada por uma grande heterogeneidade de mutações no gene que codifica o fator VIII da coagulação sanguínea, apresenta manifestações graves causadas por inversões gênicas.

Mutação reversa
Segunda mutação que ocorre em uma mesma posição do gene em que houve uma mutação prévia, restaurando a sequência gênica original

Mutação supressora
Segunda mutação em um local diferente no genoma que compensa os efeitos da primeira mutação

Embora seja um evento raro, mutações em um segundo sítio podem reverter o efeito nocivo de uma mutação previamente ocorrida, restaurando o fenótipo normal. A esse mecanismo dá-se o nome de mutação reversa. Assim, uma mutação pontual pode ser revertida por uma segunda mutação pontual, uma deleção pode ser neutralizada por uma inserção do material deletado e uma inserção pode ser revertida por uma deleção do material inserido. Além disso, ainda podem ocorrer mutações em outro(s) gene(s) que podem suprimir o(s) efeito(s) da primeira mutação no gene original, o que é nomeado de mutação supressora.

Figura 3.4 Inversão gênica e seus efeitos. (A) Excisão de três bases seguida de sua reinserção em orientação inversa causando a troca de um aminoácido; (B) inversão de quatro bases destruindo um códon de parada, o que faz com que aminoácidos adicionais sejam codificados até que um códon de término seja alcançado. O efeito inverso também pode ocorrer (criação de um códon de parada prematuro).

■ Mutações em íntrons

Durante muito tempo, acreditou-se que, pelo fato de os íntrons serem regiões não codificantes excisadas do RNA durante o seu processamento, qualquer mutação que ocorresse nessas regiões não repercutiria na formação da cadeia polipeptídica original e, portanto, não ocasionaria um fenótipo anormal.

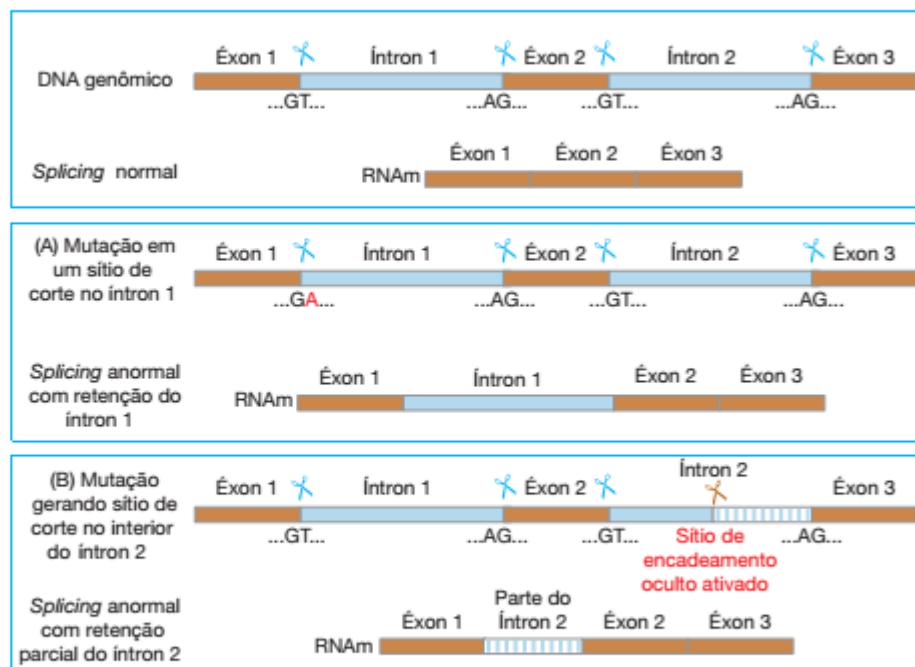
Assim, é crescente o número de distúrbios genéticos associados à presença de mutações em íntrons, incluindo vários tipos de câncer, doença de Alzheimer, deficiência intelectual, ataxias, dentre outras.

Como visto no Capítulo 2, o mecanismo de *splicing* pelo qual os íntrons são removidos do RNA primário e os éxons são reunidos para formar o mRNA maduro depende essencialmente do reconhecimento de sequências consenso localizadas nos limites entre um íntron e os éxons adjacentes, bem como da participação de sequências reguladoras presentes no interior dos íntrons, que são capazes de modular todo o processo. Assim, mutações que aconteçam nos sítios conservados de clivagem ou nas sequências intrônicas reguladoras podem ocasionar um encadeamento aberrante, em consequência da retenção de sequências intrônicas na molécula de mRNA maduro. Como resultado, ocorre introdução de códons adicionais na molécula de RNA, podendo inclusive ocorrer mudança na fase de leitura com subsequente inserção de aminoácidos extras na cadeia polipeptídica. Além disso, existem sequências no interior dos íntrons que, embora não estejam envolvidas no mecanismo de *splicing*, apresentam alguma similaridade de sequência com aquelas utilizadas como sítios de corte (*sítios de encadeamento ocultos*). Mutações nessas regiões podem transformar essas sequências em sítios de reconhecimento consensuais, também ocasionando um encadeamento anormal (Figura 3.5).

■ Mutação dinâmica | Amplificação de sequências polimórficas instáveis

O genoma humano apresenta significativa redundância, visto que 14% (420 Mb) dos 3 bilhões de pares de bases que o compõem são atribuídos a sequências de DNA repetitivo.

Figura 3.5 Efeitos de mutações intrônicas que alteram o *splicing*. (A) Substituição de uma base (T→A) em um sítio conversado utilizado para clivagem do íntron causando retenção completa do íntron 1 no mRNA maduro. (B) Uma mutação em uma sequência no interior do íntron 2 que se assemelha a uma sequência de corte de *splicing* pode transformá-la em uma sequência de clivagem ativa, causando retenção da parte final do íntron. Em ambos os casos, as consequências podem ser mudança de fase de leitura, criação de códon de parada prematuro ou introdução de aminoácidos adicionais na proteína.

Microssatélites
Sequências de DNA que se repetem consecutivamente e contêm número de repetições variável entre diferentes indivíduos

As sequências de DNA que se repetem consecutivamente (*em tandem*) em nosso genoma podem ser classificadas de acordo com o número de bases em cada unidade de repetição e a extensão dessas repetições (Capítulo 1). Dentre as diferentes classes de repetições em "tandem" do DNA, podemos encontrar os microssatélites, os quais contêm unidades de repetição variando de 1 pb a 4 pb, dispostas em blocos que em geral não ultrapassam 150 pb em comprimento. Os microssatélites representam a porção mais significante da família de sequências repetitivas de nosso genoma e incluem as *repetições trinucleotídicas* que são ocasionalmente associadas a doenças, quando expandidas de maneira anormal com o passar das gerações.

Mutação dinâmica
Expansão instável de microssatélites com o passar das gerações

Os microssatélites são normalmente polimórficos na população, o que significa dizer que o número de repetições que um indivíduo "A" apresenta não é necessariamente igual ao de um indivíduo "B", dentro de um limiar de normalidade. Contudo, o número dessas repetições pode aumentar com o passar das gerações, tornando-se instáveis e a esse mecanismo mutacional é dado o nome de mutação dinâmica. Esse tipo de mutação foi identificado na década de 1990 e, de lá para cá, várias doenças genéticas que têm como causa a instabilidade de repetições trinucleotídicas foram identificadas (síndrome do X frágil, deficiência intelectual FRAXE, distrofia miotônica, ataxia de Freidreich) (Tabela 3.1). Posteriormente, doenças associadas a expansões de tetranucleotídios (distrofia miotônica tipo 2), pentanucleotídios (ataxia espinocerebelar tipo 10) e até uma doença causada por uma repetição de uma sequência formada por 12 nucleotídios (epilepsia mioclônica juvenil) foram caracterizadas.

As sequências repetitivas propensas à expansão podem ser encontradas em regiões codificadoras, íntrons e regiões 3′ e 5′ não traduzidas e podem causar perda e/ou ganho de função gênica.

Nos casos em que acontece perda de função, o produto gênico deixa de ser produzido, ao passo que quando ocorre ganho funcional, geralmente associado a expansões de nucleotídios em regiões codificantes, o gene passa a produzir quantidade maior de determinado aminoácido, que muitas vezes passa a se acumular, provocando consequências nocivas para a célula, como neurotoxicidade. Então, poderíamos pensar: se as regiões 3′e 5′ não são traduzidas, como mutações nestas regiões podem dar origem a um fenótipo anormal? Da

■ **Tabela 3.1** Sumário das diferentes classes de doenças associadas a repetições trinucleotídicas.

Doença	*Locus*	Gene	Proteína	Repetição	Faixa normal	Faixa anormal	Local da repetição	Ganho/perda de função
Sítios frágeis								
Síndrome do X frágil	Xq27.3	*FMR1*	FMRP	CGG	6 a 54	60 a 200 (pré-mutação) > 200 (mutação completa)	5′ não traduzida	Perda
Deficiência intelectual FRAXE	Xq28	*FMR2*	FMR2	CCG	6 a 31	61 a 200 (pré-mutação) > 200 (mutação completa)	5′ não traduzida	Perda
Distrofia miotônica	19q13	*DMPK*	Proteinoquinase (DMPK)	CTG	5 a 37	> 50	3′ não traduzida	? Perda e ganho
CAG/doenças de poliglutaminas								
Doença de Huntington	4p16.3	*HD*	Huntingtina	CAG	6 a 35	36 a 121	Codificante (poliglutamina)	Ganho
Atrofia dentatorubral-palidoluisiana (síndrome de Haw River)	12p13.31	*DRPLA*	Atrofina-1	CAG	6 a 35	49 a 88	Codificante (poliglutamina)	Ganho
Atrofia muscular espinobulbar (doença de Kennedy)	Xq13-21	*AR*	Receptor de androgênio (AR)	CAG	9 a 36	38 a 62	Codificante (poliglutamina)	Ganho
Ataxia espinocerebelar tipo 1	6p23	*SCA1*	Ataxina-1	CAG	6 a 44	39 a 82	Codificante (poliglutamina)	Ganho
Ataxia espinocerebelar tipo 2	12q24.1	*SCA2*	Ataxina-2	CAG	15 a 31	36 a 63	Codificante (poliglutamina)	Ganho
Ataxia espinocerebelar tipo 3 (doença de Machado-Joseph)	14q23.1	*SCA3 (MJD1)*	Ataxina-3	CAG	12 a 40	55 a 84	Codificante (poliglutamina)	Ganho
Ataxia espinocerebelar tipo 6	19p13	*SCA6*	Subunidade de canal de cálcio voltagem-dependente α_{1A}	CAG	4 a 18	21 a 33	Codificante (poliglutamina)	Ganho
Ataxia espinocerebelar tipo 7	3p12-13	*SCA7*	Ataxina-7	CAG	4 a 35	37- 306	Codificante (poliglutamina)	Ganho
Ataxia espinocerebelar tipo 8	13q21	*SCA8*	?	CTG	16 a 37	110-<250?	Éxon terminal 3′ (antissenso?)	? Perda ou ganho
Ataxia espinocerebelar tipo 12	5q31-33	*SCA12*	PP2A-PR55β (PPP2R2B)	CAG	7 a 28	66 a 78	5′ não traduzida	? Perda
Ataxia espinocerebelar tipo 17	6q27	*SCA17*	Proteína de ligação ao TATA Box (TBP)	CAG	29 a 42	47 a 55	Codificante (poliglutamina)	? Perda ou ganho
Ataxia de Friedreich	9q13-21.1	*X25*	Frataxina	GAA	7 a 34	34 a 80 (pré-mutação) > 100 (mutação completa)	Íntron	Perda
Doenças de polialaninas								
Simpolidactilia	2q31-q32	*HOXD13*	HOXD13	GCG/A/T	15	22 a 29	Codificante (polialanina)	Ganho
Displasia cleidocranial	6p21	*CBFA1*	Proteína de ligação a elemento cis-atuante	GCG/GCT	17	27	Codificante (polialanina)	? Perda e ganho
Distrofia muscular oculofaríngea	14q11.2-q13	*PABPE*	Proteína de ligação à poli-A	GCG	6	8 a 13	Codificante (polialanina)	? Ganho

mesma maneira como abordamos as mutações em íntrons, embora não sejam traduzidas, as regiões 3′ e 5′ são transcritas, servindo como elementos imprescindíveis para a adição da cauda poli A e para o capeamento do mRNA, respectivamente, evitando a degradação no trajeto do núcleo para o citoplasma. Além disso, a região 5′ não traduzida também funciona como sítio de ligação de fatores proteicos, visando à regulação da expressão gênica. Isso explica todas as consequências deletérias associadas à amplificação anormal de unidades de repetições em tais regiões.

Acredita-se que, em virtude da alta homologia entre as unidades de repetição, a explicação mais provável para o surgimento de instabilidade nessas regiões de microssatélites seria o pareamento incorreto de repetições localizadas em cromossomos homólogos por *deslize das fitas*. Esse mecanismo leva à formação de estruturas secundárias em forma de grampo de cabelo, chamadas de *hairpin*, fazendo com que unidades de repetição pareiem erroneamente com outras repetições sucessivas na mesma fita de DNA, prejudicando a progressão da DNA polimerase durante a replicação. Caso a estrutura *hairpin* seja formada na fita molde, a fita nascente pode conter deleção em suas unidades de repetição e, caso as alças secundárias ocorram na fita nascente, esta apresentará unidades de repetição adicionais, levando à instabilidade. O curioso é que a inserção de unidades de repetição extras é muito mais frequente que as deleções, ocasionando a expansão sucessiva das unidades de repetição com o passar das gerações (Figura 3.6).

A expansão das unidades de repetições também depende da origem parental.

Em algumas doenças, a expansão da sequência de nucleotídios ocorre apenas se aquele alelo for herdado por via materna, como na síndrome do X frágil; em outras, apenas por via paterna, como na doença de Huntington, e existe ainda um subgrupo de doenças, nas quais a possibilidade de expansão independe da origem parental, como no deficiência intelectual FRAXE. Muitas vezes, o aumento sucessivo do número de repetições ao longo das gerações está relacionado a uma apresentação clínica cada vez mais grave e precoce da doença, fenômeno ao qual é dado o nome de antecipação genética.

Antecipação genética
Aumento da severidade da doença e início progressivamente mais precoce ao longo de sucessivas gerações

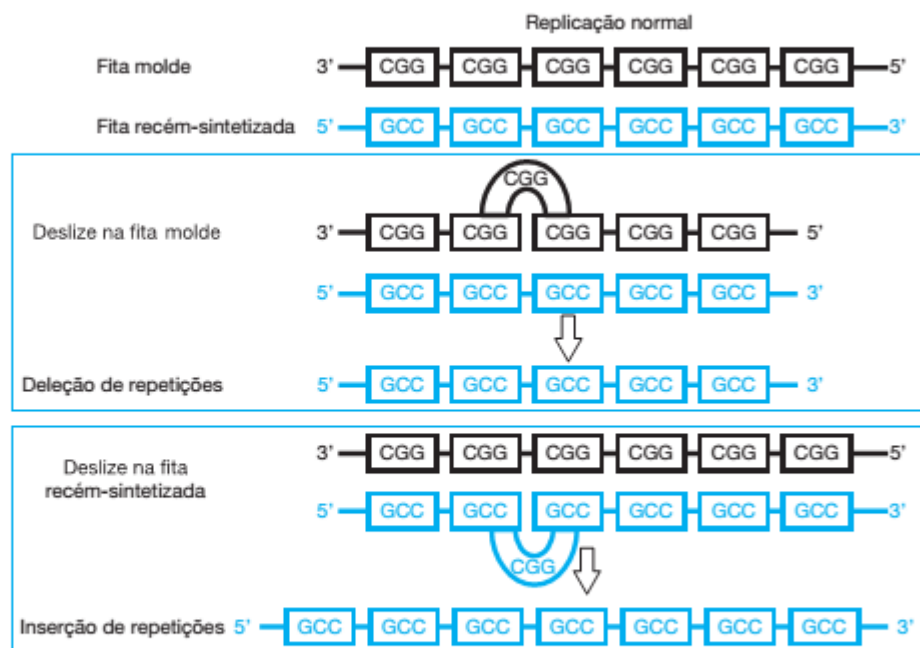



Figura 3.6 Mecanismo de pareamento incorreto por deslize das fitas com formação de estruturas secundárias, ocasionando deleção ou expansão de microssatélites na fita recém-sintetizada.

G·F Genética em Foco

Variações no número de cópias gênicas

Recentemente, as variações no número de cópias gênicas (CNV, do inglês *copy number variation*), representadas essencialmente por microdeleções e microduplicações gênicas submicroscópicas maiores que 1 kb, têm sido consideradas o tipo de variação estrutural mais prevalente no genoma humano. Entre dois indivíduos, o número de diferenças causadas pelas CNV chega a ser 100 vezes maior que a atribuída aos SNP, tendo um grande impacto na diversidade populacional. No entanto, o envolvimento de genes importantes e de outros elementos regulatórios englobados por essas regiões variáveis deixam evidente que, em determinadas situações, as CNV apresentam consequências funcionais, contribuindo para um número crescente de doenças, tais como câncer, doenças cardiovasculares, suscetibilidade ao HIV, distrofia muscular, deficiência intelectual, doenças de Alzheimer e de Parkinson, doenças metabólicas, autismo e esquizofrenia.

As CNV podem afetar a expressão de um ou múltiplos genes por diferentes caminhos. Quando um gene inteiro está envolvido em um rearranjo, espera-se como consequência simples um efeito de dosagem, pelo qual microdeleções ou microduplicações neste gene levam a diminuição ou aumento dos níveis de expressão gênica, respectivamente. No entanto, nem toda CNV resultará em mudanças nos níveis de expressão gênica. Parte dos efeitos das CNV pode ser causada pela disrrupção das sequências codificantes de um dado gene, representadas, por exemplo, pela remoção de parte dos éxons ou pela alteração de elementos reguladores e outras regiões funcionais. Devemos lembrar, ainda, que a perda ou ganho de material exônico também podem resultar em mutações *frameshift* e novos genes fusionados podem ser produzidos. Contudo, microdeleções e microduplicações fora da região codificante talvez afetem a expressão gênica pela mudança de localização ou de eficiência de importantes elementos reguladores, como os efeitos de posição. Deleções em um gene que funciona como elemento repressor à montante pode levar a um aumento dos níveis transcricionais de um gene por ele regulado, enquanto duplicações de sequências à jusante do promotor talvez diminuam os níveis de expressão, devido à alteração da posição relativa do promotor em relação ao gene. Uma causa comum para a formação de CNV é a recombinação homóloga não alélica entre sequências repetidas de DNA com baixo número de cópias.

O advento de novas metodologias genômicas capazes de identificar CNV e caracterizar as regiões críticas mínimas compartilhadas por diferentes indivíduos apresentando uma determinada doença, certamente, trará grandes benefícios para a elaboração de novas estratégias preventivas e terapêuticas. Além disso, a caracterização dos mecanismos moleculares envolvidos na formação de CNV é uma ferramenta importante na compreensão da evolução do genoma humano, aprimorando nosso conhecimento sobre a variação genética de nossa espécie.

■ Mutações no genoma mitocondrial

Não podemos esquecer que, da mesma maneira que o genoma nuclear, o material genético contido nas mitocôndrias também pode sofrer mutações. O papel da mitocôndria na célula é produzir energia mediante a fosforilação oxidativa, uma reação em que os nutrientes orgânicos são oxidados pelo oxigênio molecular e a energia química liberada é usada para gerar ATP. O DNA mitocondrial se replica mais que o genoma nuclear, principalmente naqueles tecidos em que há grande gasto de energia, como o muscular. A elevada frequência de produção de

intermediários reativos do oxigênio pela cadeia respiratória pode causar um substancial estrago oxidativo no DNA mitocondrial. O fato de o genoma mitocondrial apresentar sistemas de reparo menos efetivos que aqueles presentes no DNA nuclear e não estar protegido por histonas faz com que se torne mais suscetível à ocorrência de mutações.

Há dois tipos de mutações no genoma mitocondrial: a *direta*, que afeta um gene do genoma mitocondrial, e a *indireta*, na qual ocorre alteração de um gene nuclear codificante de um componente essencial ao metabolismo mitocondrial.

Muitas doenças associadas a mutações mitocondriais já foram descritas, incluindo a neuropatia ótica de Leber, epilepsia mioclônica, miopatia mitocondrial com cardiopatia, encefalopatia mitocondrial, fraqueza muscular neurogênica com ataxia e oftalmologia externa progressiva. O que podemos observar é que essas doenças têm uma ampla variabilidade na apresentação clínica, até mesmo para uma mesma mutação. Por que isso ocorre? Em primeiro lugar, devemos considerar que dentro de uma célula temos várias mitocôndrias e dentro de uma mesma mitocôndria temos várias moléculas circulares de DNA. Dessa maneira, a apresentação fenotípica de uma mutação mitocondrial direta dependerá da proporção entre o número de moléculas de DNA mitocondrial normais e o número de mitocôndrias mutadas. À coexistência de mitocôndrias com sequências de DNA diferentes em uma mesma célula ou tecido, é dado o nome de heteroplasmia. Outro fator que pode modular a variabilidade clínica das doenças associadas a mutações mitocondriais é a presença de genes nucleares modificadores. Como dispomos de *backgrounds* genéticos distintos, indivíduos portadores de um mesmo tipo de mutação mitocondrial, mesmo em homoplasmia (uniformidade quanto à sequência de DNA em todas as mitocôndrias), podem apresentar fenótipos diferentes causados por uma interação diferenciada de seus genomas nucleares com o genoma mitocondrial alterado.

Heteroplasmia
Presença de, pelo menos, dois tipos de sequências de DNA mitocondrial em um indivíduo

Homoplasmia
Presença de somente um tipo de sequência de DNA mitocondrial em um indivíduo

Outro ponto interessante é o fato de a mitocôndria ter origem bacteriana, tendo se instalado na célula humana há milhares de anos por *endossimbiose*. Portanto, o uso intenso de antibióticos, especificamente destinados a atacarem o DNA bacteriano durante a sua replicação, pode causar alterações no genoma mitocondrial e contribuir para o surgimento de doenças, incluindo a surdez.

Nomenclatura das mutações

Existem regras para citações de variantes gênicas, estabelecidas pela *Human Genome Variation Society*, uma maneira uniforme e eficaz de os pesquisadores se comunicarem entre si. Pela nomenclatura estabelecida, podemos descrever uma alteração surgida na sequência de DNA em diferentes níveis (DNA, RNA e proteína). Em geral, a descrição de uma variante é precedida de uma letra, que corresponde ao nível em que estamos nos referindo, seguida de um número que representa a localização da alteração e a mudança ocorrida. Os diferentes níveis são ilustrados pelas seguintes abreviações: DNA ("*g*" para sequência *g*enômica ou "*c*" para uma sequência de DNA *c*omplementar e "*m*" para sequência *m*itocondrial), RNA ("*r*" para sequência de *R*NA) e proteína ("*p*" para sequência de *p*roteína). Abreviações/símbolos ">", "del", "ins" ou "inv" são utilizados para descrever substituições, deleções, inserções e inversões, respectivamente.

O DNA complementar é numerado a partir do número "1", que representa a "A" do códon de início da tradução ATG, contrariamente ao DNA genômico, no qual a numeração dos nucleotídios começa com 1 no primeiro nucleotídio do arquivo de referência do banco de dados. Em ambos os casos, os nucleotídios alterados devem aparecer em letra maiúscula. Uma variante c.677C>T, por exemplo, significa dizer que na posição 677 de determinado DNA complementar houve a substituição de uma citosina por uma guanina. Quando a alteração envolve mais de uma base, os resíduos afetados são indicados por "_" entre o primeiro e o último nucleotídio afetados (c.76_78delACT). No caso de variantes intrônicas, a alteração é numerada de acordo com sua localização no íntron em relação aos éxons adjacentes. Caso a variante apareça no início do íntron, é utilizado o número do último nucleotídio do éxon precedente, seguido de um sinal "+" e a posição no íntron (c.77+1G). Quando está no final do íntron, utiliza-se o número do primeiro nucleotídio do éxon seguinte, seguido de um "-" e a posição retroativa no íntron (c.78-1G). Para o DNA genômico, os sinais de "+" ou "-" não são usados.

Em nível de RNA, as regras são basicamente as mesmas utilizadas para as sequências de DNA com a diferença de que os nucleotídios afetados devem vir em letra minúscula (ex r.76a>u). Particularidades são ainda utilizadas quando o gene tiver mais de um transcrito decorrente da recomposição alternativa dos éxons.

Para proteínas, a descrição da localização da variante é feita com base na numeração, iniciada da metionina gerada pelo códon de início da tradução, que recebe o número 1. Aminoácidos envolvidos na alteração devem ser descritos pelas três letras que o representam no código genético universal, sempre começando com a primeira letra maiúscula (p.Lys76Asn). Podemos observar que, nesse caso, não se costuma utilizar o símbolo ">". Quando uma alteração faz com que um aminoácido seja substituído por um códon de parada prematuro, utiliza-se o símbolo "X" (p.Trp26X).

■ Agentes mutagênicos

Conforme abordado, as mutações podem ser induzidas por *agentes químicos e/ou físicos*, encontrados nos meios intra e extracelular, que aumentam a taxa de ocorrência da mutação. A partir de agora, discutiremos um pouco mais a ação desses fatores mutagênicos na introdução de alterações na sequência de DNA, quando estas escapam da ação dos mecanismos de reparo.

Os *agentes físicos* abrangem ampla gama de fatores que incluem o calor, a luz ultravioleta e as radiações ionizantes. O *calor* pode clivar as ligações entre as bases nitrogenadas e seus açúcares, formando sítios apurínicos ou apirimidínicos (sem purinas ou sem pirimidinas) na molécula de DNA. Os erros não reparados podem causar mutações de ponto ou deleções. A *luz ultravioleta* é uma radiação não ionizante, que penetra apenas na camada superficial de nossas células e, como apresenta baixa energia, não é capaz de causar ionização. No entanto, a luz ultravioleta é fortemente absorvida por pirimidinas e purinas, o que pode resultar na formação, principalmente, de dímeros entre pirimidinas adjacentes na molécula de DNA. Esse processo de dimerização faz com que duas bases adjacentes se liguem entre elas em um mesmo filamento de DNA, o que comumente acontece entre duas timinas subsequentes. Quando não reparada, essa alteração pode introduzir uma deleção na molécula de DNA. Alternativamente, a mudança na configuração da molécula de DNA contendo o dímero pode inibir a replicação e a transcrição (Figura 3.7).

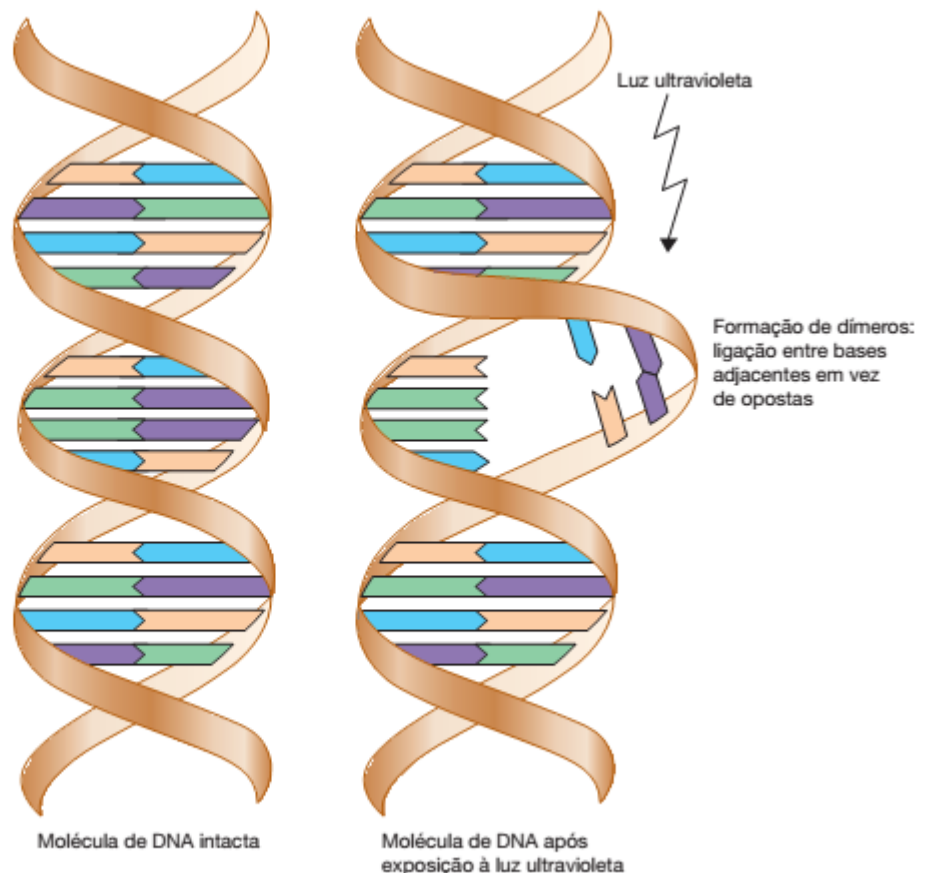



Figura 3.7 Formação de dímeros de pirimidinas após a exposição pela luz ultravioleta.

Ao contrário da luz ultravioleta, as *radiações ionizantes*, representadas pelos raios X, α, β e γ, têm energia suficiente para atravessarem as células e promoverem a liberação de elétrons, produzindo íons. Os íons colidem com outras moléculas, resultando na liberação de mais elétrons. A instabilidade ocasionada pode alterar a estrutura das bases e romper as ligações açúcar-fosfato, ocasionando quebras cromossômicas e outras formas de rearranjos gênicos. Em geral, a frequência de mutações causadas por radiações ionizantes é diretamente proporcional à dose de irradiação e ao tempo de exposição.

Os *agentes químicos* podem induzir danos à molécula de DNA de diferentes maneiras. Substâncias químicas que apresentam similaridade estrutural com as bases padrões do DNA (*análogos de base*) podem ser incorporadas erroneamente na molécula de DNA durante a sua replicação, visto que a DNA polimerase não consegue distingui-las das bases normais, induzindo mutações, em geral substituições. Um exemplo é o 5-bromouracil, que apresenta semelhança estrutural com a timina, com a diferença de um átomo de bromo (-Br) no lugar do grupamento metil (CH_3). O 5-bromouracil pode ser incorporado no lugar da timina durante a replicação, pareando com a adenina. Além disso, pode sofrer ainda mudança tautomérica em sua estrutura, passando da forma cetônica para a forma enólica. Nesse novo formato, o 5-bromouracil pode parear com a guanina, o que leva à substituição do par original A-T por C-G na molécula mutada (Figura 3.8). Outro agente químico indutor de mutações é o 2-aminopurina, que é análogo da adenina e, em geral, pareia com a timina. No entanto, o 2-aminopurina pode parear de maneira errônea com a citosina, provocando uma transição do par T-A por C-G durante a replicação.

Os *agentes intercalantes* (alaranjado de acridina, proflavina, dioxina e brometo de etídio) representam mais um modelo de indutores químicos de mutações. Essas substâncias têm a capacidade de se intercalarem na molécula de DNA, propriedade esta que faz com que sejam amplamente utilizados nos laboratórios de Biologia Molecular para análise do DNA. O mais conhecido desses agentes é o brometo de etídio, um corante composto por quatro anéis aromáticos, o qual se assemelha em tamanho a um par de bases purina-pirimidina. Com isso, o brometo de etídio pode ser incorporado de forma intercalada à dupla fita, causando deleções e inserções de um ou mais pares de bases, podendo inclusive alterar a matriz de leitura gênica.

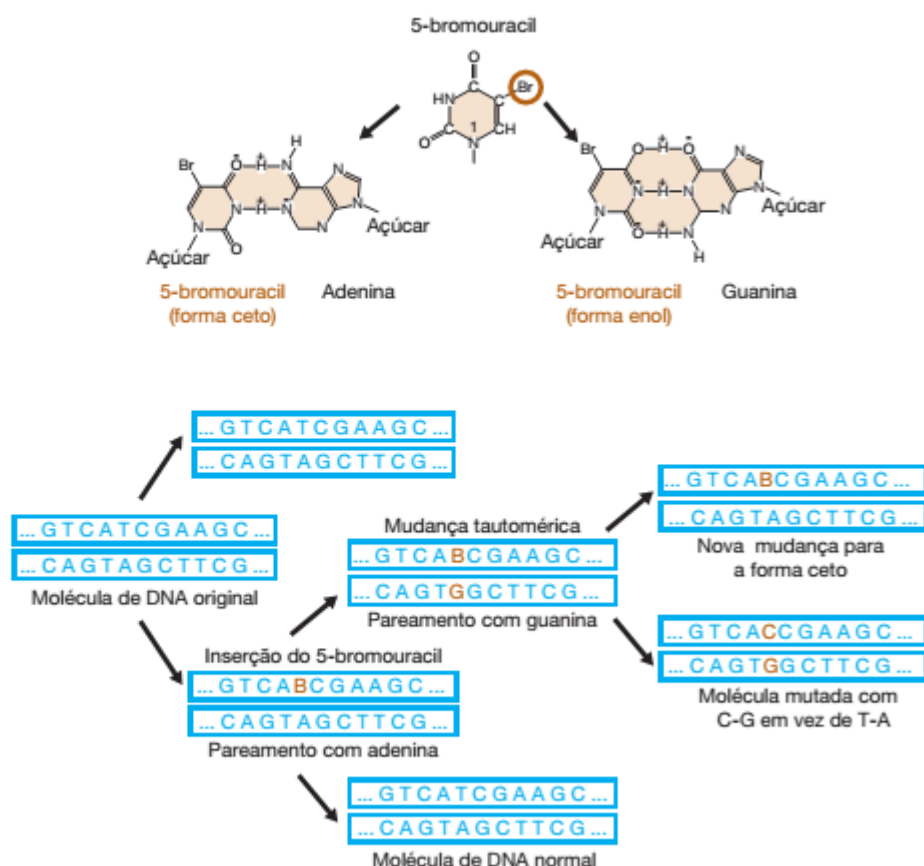



Figura 3.8 Efeitos mutagênicos do 5-bromouracil.

Agentes alquilantes (metil metano sulfonato e etil metano sulfonato) interagem quimicamente com os centros nucleofílicos do DNA, adicionando grupamentos alquil. Exemplos adicionais de potenciais agentes mutagênicos químicos incluem as espécies reativas do oxigênio (radicais superóxido, peróxido de hidrogênio), produzidos no metabolismo aeróbico normal, e os agentes desaminantes (presentes em conservantes de carne), os quais interagem com as bases que têm o radical amino, causando desaminação. A indução de mutações também pode ocorrer pela movimentação de elementos genéticos móveis, os chamados transpósons, os quais são capazes de se inserir em uma sequência gênica, podendo torná-la não funcional.

■ Mecanismos de reparo

Grande parte das alterações que acontecem em nosso genoma não chega a se manifestar fenotipicamente. Além dos diferentes mecanismos utilizados por uma célula ou organismo para evitar a propagação de uma mutação somática ou germinativa, como a inviabilidade celular e a diminuição da capacidade reprodutiva de um indivíduo, nosso genoma pode contar ainda com a precisa atuação de diferentes mecanismos de reparo. Em virtude da diversidade de mecanismos de reparo existentes desde bactérias a humanos, torna-se evidente a necessidade de manutenção da integridade do genoma, evitando a perpetuação da maioria das mutações. Todas as nossas células dispõem de caminhos alternativos de reparo de lesões na molécula de DNA. A escolha por cada um desses caminhos dependerá do tipo de dano e da situação. A partir de agora, abordaremos os principais mecanismos de reparo identificados em nossa espécie.

Prevenção de erros

Antes mesmo que a mutação aconteça, nosso organismo dispõe de alguns sistemas enzimáticos que atuam sobre compostos potencialmente mutagênicos, neutralizando-os de forma prévia a sua interação com a dupla fita. As enzimas superóxido dismutase e catalase são ótimos exemplos de sistemas de defesa antioxidantes, tendo um papel essencial na proteção do DNA contra danos oxidativos provocados por espécies reativas do oxigênio (radicais livres). A enzima superoxido dismutase catalisa a conversão do ânion reativo superóxido ($O_2^{-\bullet}$) em peróxido de hidrogênio (H_2O_2), pela reação $2O_2^{-\bullet} + 2H^+ \rightarrow O_2 + H_2O_2$, ao passo que a enzima catalase converte o peróxido de hidrogênio em oxigênio molecular e água ($2\ H_2O_2 \rightarrow 2H_2O + O_2$), neutralizando os efeitos que os radicais livres poderiam causar à molécula de DNA.

Reparo por excisão de nucleotídios

O mecanismo de *reparo por excisão de nucleotídios* conta com a participação de quatro classes de enzimas essenciais: endonucleases, exonucleases, DNA polimerases e DNA ligases (Figura 3.9). Primeiramente, as endonucleases de reparo reconhecem o fragmento danificado, o qual é posteriormente excisado por exonucleases com base na quebra da ligação fosfodiéster em regiões que flanqueam a lesão, dando origem a um fragmento que tem em torno de 27 a 29 nucleotídios. Em seguida, uma DNA polimerase (ε e δ em humanos) utiliza-se do filamento de DNA complementar não danificado para preencher a lacuna deixada no processo de excisão, usando-o como molde. Por último, uma DNA ligase completa o processo, religando o novo fragmento ao restante da molécula. O sistema de reparo por excisão de nucleotídios é utilizado para reparos considerados grandes.

Reparo por excisão de base

Muitas vezes, o dano ocasionado à molécula de DNA não provoca distorção suficientemente grande para ser inicialmente reconhecida pelo sistema de reparo por excisão de nucleotídios, como nos casos de alterações na estrutura química das bases. Nesses casos, o início do reparo conta com a participação de enzimas DNA glicosilases que reconhecem a(s) base(s) alterada(s), clivando as ligações glicosídicas entre a base anormal e o açúcar. Existem tipos específicos de DNA glicosilases para cada uma das alterações encontradas (bases desaminadas, oxidadas etc.),

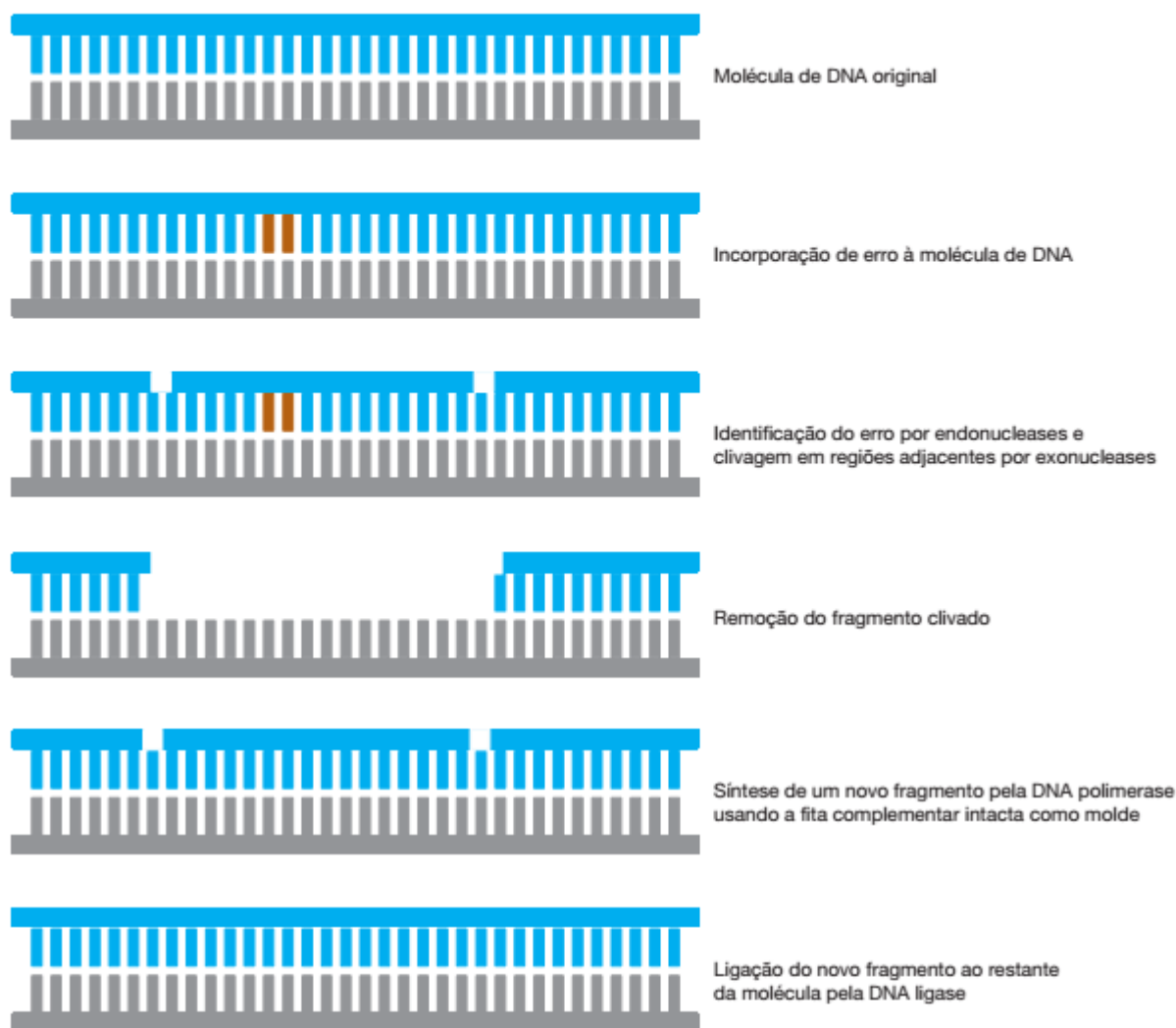



Figura 3.9 Mecanismo de reparo do DNA por excisão de nucleotídios.

mas, ao clivarem as ligações glicosídicas, todas elas resultam em sítios apurínicos ou apirimidínicos (sítios AP), os quais são reconhecidos por endonucleases AP, as quais atuam juntamente com fosfodiesterases, excisando os grupos açúcar-fosfato naqueles locais onde as bases estão ausentes. Em seguida, uma DNA polimerase substitui o nucleotídio ausente, utilizando como molde a fita complementar não danificada e uma DNA ligase restaura a fita, selando as extremidades da clivagem. Um exemplo comum de reparo por excisão de base é o que acontece na correção da desaminação da citosina à uracila por meio de uma uracila-DNA-glicosilase.

Reparo de mau pareamento

Algumas bases mal pareadas conseguem escapar da atividade exonucleásica 3′→5′ de revisão realizada pela DNA polimerase durante a replicação. O mecanismo de *reparo de mau pareamento* busca esses pareamentos incorretos que permaneceram entre as bases que formam a molécula de DNA após sua replicação, como o pareamento de um T com um G, por exemplo. Como se tratam de bases padrões normalmente encontradas na molécula de DNA, para que a maquinaria de reparo por mau pareamento reconheça o nucleotídio do par incorporado de maneira inapropriada, o sistema tem de distinguir a fita antiga daquela recém-sintetizada, corrigindo sempre a sequência presente nesta última. Caso isso não ocorra, a remoção de uma das bases será aleatória, havendo a probabilidade de 50% de a

base correta ser removida e a mutação poder ser propagada às células-filha. No entanto, a distinção entre os dois filamentos (molde e recém-sintetizado) pode ser feita com base no padrão de metilação. Embora a fita recém-sintetizada seja cópia da fita molde e os padrões de metilação sejam mantidos, existe um período entre a replicação e a adição de grupamentos metil às citosinas específicas na fita nova pela ação de DNA metiltransferases. Enquanto essa adição não acontece, o sistema de reparo de mau pareamento identifica a fita não metilada como sendo a fita recém-sintetizada, corrigindo-a com base na sequência disposta na fita antiga.

Reparo por recombinação

Esse sistema de reparo recebe esse nome pelo fato de sua atividade fazer uso de pro-priedades inerentes ao processo de recombinação do DNA, sendo utilizado de maneira efetiva nos casos em que um dos filamentos da molécula de DNA mãe contém bases danificadas. Nessas situações, o dano pode causar distorções estruturais em uma das fitas do duplex, como a provocada por dímeros de pirimidinas. Quando o DNA é replicado, a maquinaria de replicação passa pelo ponto da lesão sem copiá-la, resultando na formação de uma lacuna de tamanho considerável na fita recém-sintetizada do filamento molde danificado. A informação contida no sítio danificado é perdida e somente pode ser recuperada pela utilização de outra molécula idêntica de DNA, conforme veremos a seguir.

Devemos notar que as duas moléculas de fita dupla sintetizadas a partir da molécula molde são estruturalmente distintas. Uma delas contém o filamento parental contendo a lesão e um filamento complementar apresentando a lacuna, enquanto a outra molécula contém uma fita parental não danificada, que foi copiada em uma fita recém-sintetizada complementar, também normal. Em seguida, o filamento de DNA similar presente na molécula intacta é utilizado para preencher a lacuna presente no filamento recém-sintetizado da primeira molécula. Com isso, a molécula-filha anteriormente lesada em ambas as fitas (molécula receptora) passa a apresentar o dano apenas na fita parental, enquanto a molécula doadora previamente intacta passa a exibir agora uma lacuna em uma de suas fitas. Desde então, essa lacuna pode ser preenchida pela síntese de reparo, dando origem novamente a uma molécula normal. Desse modo, apenas a fita originalmente mutada permanecerá com a alteração e, até que esta seja removida pelos sistemas de reparo de excisão, novos ciclos de replicação envolverão essa forma de reparo por recombinação (Figura 3.10).

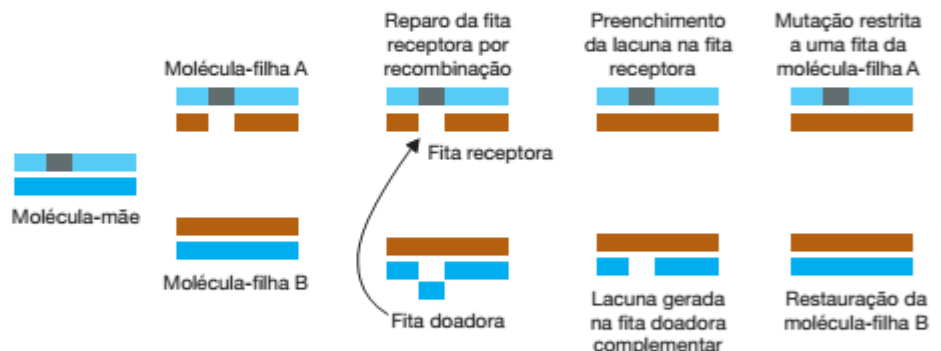



Figura 3.10 Reparo por recombinação. Após o surgimento de uma mutação em uma das fitas da molécula parental, a maquinaria de replicação cria uma lacuna na fita recém-sintetizada gerada com base na fita que tem a lesão na molécula-filha A, enquanto a molécula-filha B é estruturalmente normal. A lacuna é preenchida por recombinação com base na fita idêntica, presente na segunda molécula-filha que, por sua vez, passa a exibir um intervalo. Este é preenchido pelos sistemas de reparo, fazendo com que a mutação fique confinada a apenas uma das 4 fitas envolvidas, e todo o processo é repetido até que o dano seja removido por excisão ou a célula morra.

Mutações em genes participantes dos mecanismos de reparo

O controle do ciclo celular é exercido pela atuação de *check points* (Capítulo 2). Em geral, quando uma mutação acontece, as células lesadas sofrem retardo de seu ciclo celular, de modo a possibilitar que o dano seja reparado ou que a célula seja induzida a sofrer apoptose, impedindo a perpetuação da mutação para as células-filha. Isso é feito com a participação da proteína ATM, uma serina-treonina quinase que identifica a lesão e libera um sinal para a proteína p53, considerada a guardiã de nosso genoma. A p53 pode ativar proteínas envolvidas no reparo do DNA, induzir a parada no ciclo celular até que a lesão seja corrigida ou induzir a apoptose, quando o erro no DNA é irreparável.

A imprescindibilidade da complexa rede de sinalização celular e dos vários mecanismos de reparo na eliminação de erros na replicação do DNA evidencia-se ainda mais ao analisarmos doenças causadas por mutações em genes desses sistemas.

A ataxia telangiectasia é um distúrbio causado por mutações no gene que codifica a proteína ATM. Nos indivíduos afetados pela doença, a deficiência da proteína faz com que células lesadas permaneçam no ciclo celular sem que haja a checagem pelos pontos de controle. Além da alta predisposição ao câncer, os portadores dessa doença apresentam pouco equilíbrio e coordenação (ataxia), marcas vermelhas na face (telangiectasia), deficiência imune e retardo na maturação sexual.

Pessoas com xeroderma pigmentoso apresentam alterações em genes envolvidos no reparo de danos ao DNA induzidos pela luz ultravioleta. Dessa maneira, os portadores da doença não conseguem reparar alterações como dímeros de timinas, tornando-os extremamente sensíveis à luz solar e ocasionando um risco aumentado de câncer de pele de início precoce, além de distúrbios neurológicos provocados provavelmente por morte prematura de células nervosas. Indivíduos com xeroderma pigmentoso devem viver em ambientes iluminados por luz artificial e fazer uso rotineiro de bloqueadores solares, bem como sair de casa apenas durante a noite. Qualquer exposição ao sol pode causar o aparecimento de lesões degenerativas na pele. Existem diferentes genes envolvidos no sistema de reparo de excisão de nucleotídios, cujas mutações estão associadas ao xeroderma pigmentoso, entre eles os genes *XPA, XPB, XPC, XPD, XPE, XPF* e *XPG*. Mutações em ambos os alelos de um desses genes devem estar presentes para que o xeroderma pigmentoso se manifeste. Além disso, em um percentual dos casos, o reparo por excisão de nucleotídios é normal, mas ocorrem mutações no gene *XPV*, o qual codifica uma DNA-polimerase de translesão, relacionada com a manutenção da continuidade da síntese do DNA que sofreu lesão.

Dois outros distúrbios também resultam de disfunções no reparo por excisão de nucleotídios: a síndrome de Cockayne e a tricotiodistrofia. Na síndrome de Cockayne, os pacientes afetados têm excesso de sensibilidade à luz solar, deficiência intelectual grave, nanismo e envelhecimento precoce. Já na tricotiodistrofia, a principal característica observada é a deficiência de enxofre associada à presença de cabelos escamosos, unhas frágeis e pele sensível. Além disso, os indivíduos com a doença apresentam atraso no crescimento e deficiência intelectual. Entretanto, curiosamente, em ambas as doenças não há aumento na incidência de câncer. Outros exemplos de doenças associadas a defeitos nos mecanismos de reparo do DNA podem ser encontrados na Tabela 3.2.

Concluindo, embora as mutações aqui abordadas aconteçam durante a replicação do DNA, devemos notar que as consequências deletérias, caso não haja reparo, serão percebidas em

■ **Tabela 3.2** Exemplos de distúrbios associados a erros nos mecanismos de reparo do DNA.

Doença	Gene alterado	Causas	Características clínicas
Anemia de Fanconi	*FA*	Reparo de excisão deficiente	Anemia aplásica e anomalias congênitas
Ataxia telangiectasia	*ATM*	Deficiência na checagem do ciclo celular para reparo do DNA	Ataxia, telangiectasia, linfoma, deficiência imune
Síndrome de Bloom	*BLM*	Mutações no gene de helicase envolvido na replicação e recombinação do DNA	Baixa estatura, fotossensibilidade cutânea, aumento do risco de câncer
Síndrome de Cockayne	*CSA, CSB*	Reparo de excisão deficiente para danos provocados por luz ultravioleta	Pele fotossensível, calcificação cerebral, degeneração de substância branca, nanismo
Síndrome de Werner	*WRN*	Helicase deficiente	Baixa estatura, malignidade, envelhecimento precoce
Tricotiodistrofia	*TTDA*	Reparo de excisão deficiente	Baixa estatura, cabelo e pele anormais, prejuízo no desenvolvimento
Xeroderma pigmentoso	*XPA a XPG, XPV*	Reparo de excisão deficiente para danos provocados por luz ultravioleta	Fotossensibilidade cutânea, sardas, câncer de pele

diferentes níveis subsequentes desde o amadurecimento do RNA até as modificações pós-traducionais das proteínas. Além disso, os efeitos dessas mutações irão depender não somente do tipo, da localização e da extensão da mutação, mas também da interação dessa nova forma alélica com os produtos de todos os outros genes que compõem a malha interativa de nosso genoma. Sendo assim, embora o uso de ferramentas de Bioinformática, atrelado a ensaios experimentais modernos, tenha apresentado novas direções sobre o estudo dos efeitos patogênicos das mutações, cada indivíduo poderá expressá-las de maneira diferente.

RESUMO

- Embora o mecanismo de replicação do DNA em nosso organismo seja extremamente eficiente, em algumas situações uma mudança na sequência de nucleotídios (mutação) pode acontecer espontaneamente ou pela ação de agentes químicos ou físicos (agentes mutagênicos)
- Em nível celular, as mutações podem ser germinativas, podendo ser transmitidas aos descendentes, ou somáticas
- Em nível molecular, as mutações podem surgir de quatro tipos principais de alterações: mutações de ponto, deleções, inserções e inversões
- A maneira como a expressão do gene mutante será alterada dependerá do tipo de mutação e de sua localização ao longo de um gene
- Mutações em íntrons também podem alterar a expressão de um gene, quando acontecem em regiões conservadas envolvidas no encadeamento do RNA, levando à retenção de sequências intrônicas na molécula de mRNA maduro
- Mutações que resultam na expansão instável de microssatélites ao longo das gerações (mutação dinâmica), mesmo que presentes em regiões não traduzidas, também podem causar doenças. Algumas dessas doenças podem apresentar antecipação genética ao longo das gerações
- Mutações no genoma mitocondrial podem afetar efetivamente algum gene presente no genoma mitocondrial ou podem estar em um gene nuclear que codifica algum componente essencial para o metabolismo mitocondrial. As mutações mitocondriais podem aparecer em homoplasmia ou heteroplasmia
- Entre os agentes mutagênicos conhecidos, podemos destacar os físicos (calor, luz ultravioleta, radiação ionizante), químicos (agentes intercalantes, alquilantes e análogos de bases) e os elementos genéticos móveis
- Grande parte das alterações que acontecem em nosso genoma não chega a se manifestar em virtude, em parte, da atuação de diferentes mecanismos de reparo, entre os quais estão o reparo por excisão de nucleotídios, o reparo por excisão de base, o reparo de mau pareamento e o reparo por recombinação
- Mutações em genes participantes dos sistemas de reparo de DNA também podem ocasionar patologias.

AUTOAVALIAÇÃO

3.1 Em termos evolutivos, qual classe de mutações apresenta maior importância: as mutações somáticas ou germinativas? Por quê?

3.2 Quais consequências uma mutação pontual pode causar para o funcionamento de um gene?

3.3 Por que uma deleção de três pares de bases é menos nociva para o funcionamento de um gene do que a deleção de apenas uma base?

3.4 Como uma inversão gênica que não leva à perda de material genético pode ser nociva para a expressão de um gene?

3.5 Distinga mutação reversa de mutação supressora.

3.6 Quando uma mutação intrônica pode levar a uma alteração na funcionalidade de um gene?

3.7 Descreva como surge uma mutação dinâmica e defina o conceito de antecipação genética.

3.8 Diferencie homoplasmia de heteroplasmia.

3.9 Dê um exemplo de agente químico que induz mutações e descreva como ele age.

3.10 Em que consistem os mecanismos de reparo? Descreva um desses mecanismos.

BIBLIOGRAFIA

ALBERTS, B.; JOHNSON, A.; LEWIS, J. *et al. Molecular biology of the cell.* 4th edition, New York: Garland Science, 2002.

ALMAL, S.H.; PADM, M. Implications of gene copy-number variation in health and diseases. *Journal of Human Genetics.* 2012; 57:6-13.

BENJAMIN, L. *Genes IX.* Porto Alegre: Artmed, 2009.

CHAMARY, J.V.; HURST, L. The price of silent mutations. *Scientific American* 2009; 46-53.

dbSNP. Disponível em http://www.ncbi.nlm.nih.gov/snp. Acesso em 01/02/2012.

DEN DUNNEN, J.T.; ANTONARAKIS, S.E. Mutation nomenclature. *Curr Protoc Hum Genet.* 2003 Aug;Chapter 7:Unit 7.13.

HUMAN GENOME VARIATION SOCIETY. Disponível em: http//www.hgvs.org/. Acesso em 30/07/2011.

KONDRASHOV, F.A.; KONDRASHOV, A.S. Measurements of spontaneous rates of mutations in the recent past and the near future. *Philos Trans R Soc Lond B Biol Sci.* 2010; 365:1169-76.

MIR, L. *et al. Genômica.* São Paulo: Atheneu, 2004.

NUSSBAUM, R.L.; MCINNES, R.R.; WILLARD, H.F. *Thompson & Thompson. Genética médica.* 7ª ed., Rio de Janeiro: Elsevier, 2007.

PEARSON, C.E.; NICHOL EDAMURA, K.; CLEARY, J.D. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet.* 2005; 6:729-742.

SNUSTAD, D.P.; SIMMONS, M.J. Tradução Paulo A. Motta. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

STRACHAN, T.; READ, A.P. *Human molecular genetics.* 2nd edition. New York: Wiley-Liss, 1999.

4

Citogenética | Organização Cromossômica e Anomalias

Objetivos de estudo

- Compreender como os dois tipos de divisão celular ocorrem em nosso organismo
- Entender o processo de formação dos gametas masculinos e femininos e suas diferenças
- Compreender os mecanismos que produzem variabilidade genética dos gametas e a sua importância evolutiva
- Saber diferenciar distância física de distância genética
- Ser capaz de perceber como os cromossomos humanos podem ser identificados citogeneticamente e saber utilizar a nomenclatura internacional para a descrição de alterações cromossômicas
- Identificar os mecanismos capazes de originar as alterações cromossômicas numéricas e estruturais
- Saber reconhecer os diferentes tipos de alterações cromossômicas e suas consequências individuais ou para a progênie
- Entender as características próprias dos cromossomos sexuais, as principais alterações a eles relacionadas e suas consequências

Conceitos-chave do capítulo

Aconselhamento genético
Alelos
Anáfase meiótica
Anáfase mitótica
Análise cromossômica
Aneuploidia
Aneussomia segmentar
Anomalias cromossômicas
Anomalias cromossômicas estruturais
Anomalias cromossômicas numéricas
Autossomo
Bandeamento cromossômico
Braço p
Braço q
Cariótipo
Células diploides
Células germinativas
Células haploides
Células somáticas
Centimorgan
Centrômero
Centrossomo
Ciclo celular
Cinetócoro
Citocinese
Citogenética
Compensação de dose
Complexo sinaptonêmico
Congressão
Constrição primária
Constrição secundária
Corpúsculo polar
Cromatina
Cromossomo acrocêntrico
Cromossomo dicêntrico
Cromossomo em anel
Cromossomo Filadélfia
Cromossomos homólogos
Cromossomo marcador
Cromossomo metacêntrico
Cromossomos sexuais
Cromossomo submetacêntrico
Cromossomo telocêntrico
Cromossomo X
Cromossomo Y
Deleção
Diacinese
Diagnóstico pré-implantação
Diagnóstico pré-natal
Diplóteno
Disjunção
Distância genética
Distância física
Divisão celular
Divisão equacional
Divisão reducional
Duplicação
Espermátides
Espermatócito primário
Espermatócito secundário
Espermatogênese
Espermatozoide
Euploidia
Fase G0 do ciclo celular
Fase G1 do ciclo celular
Fase G2 do ciclo celular
Fase M do ciclo celular
Fenótipo
Fibroblastos
Fluorescence *in Situ* Hibridization (FISH)
Folato
Frequência de recombinação
Fuso mitótico
Gametas
Gametogênese
Gene MTHFR
Gene SRY
Genoma
Genótipo
Glóbulo polar
Haploinsuficiência
Haplótipo
Hemizigose
Heterocromatina constitutiva
Heterocromatina facultativa
Hermafroditismo
Hibridação *in situ* com fluorescência
Hiperdiploidia
Hiperplasia adrenal congênita
Hipodiploidia
Homens XX
Homólogos
Idade materna avançada
Imprinting genômico
Inativação do cromossomo X

Indicações para análise citogenética
Infertilidade
Inserção cromossômica
Inversão cromossômica
Intercinese
Intérfase
Inversão paracêntrica
Inversão pericêntrica
International System for Human Cytogenetic Nomenclature (ISCN)
Isocromossomo
Leptóteno
Ligação
Linhagem germinativa
Loci
Lod score
Malformações congênitas
Mapeamento físico
Mapeamento genético
Meiose
Meiose equacional
Meiose reducional
Metáfase meiótica
Metáfase mitótica
Mosaicismo
Mulheres XY
Não disjunção
Nomenclatura cromossômica
Ovocitogênese
Ovócitos primários
Ovócitos secundários
Óvulo
Origem de replicação
Origem parental
Padrão de bandas
Paquíteno
Poliploidia
Pontos de controle na mitose
Prófase
Prometáfase
Pseudo-hermafroditismo
Quadrivalente
Quiasmas
Quimeirismo
Rearranjo cromossômico
Recombinação
Recombinação desigual
Recorrência
Região pseudoautossômica
Repulsão
Segregação
Sinapse
Síndrome de Down
Síndrome de Edwards
Síndrome de genes contíguos
Síndrome de Klinefelter
Síndrome de Patau
Síndrome de Turner
Síndrome do duplo Y
Síndrome do triplo X
Telomerase
Telômero
Translocação recíproca
Translocação robertsoniana
Trissomia
Variabilidade genética
Zigóteno
Zigoto

■ Introdução

Diploide (2n)
Célula que tem os dois membros de cada par de cromossomos

Haploide (n)
Célula que tem um membro de cada par de cromossomos

Células eucarióticas somáticas
Células diploides que compõem o organismo

Linhagem germinativa
Tecido reprodutivo de um organismo

Para garantir a continuidade genética entre diferentes linhagens celulares que compõem o nosso organismo, bem como, a manutenção do número de cromossomos ao longo de sucessivas gerações, nossas células fazem uso de dois mecanismos de divisão celular: a mitose e a meiose. Embora ambos os processos tenham semelhanças estruturais em várias de suas etapas, cada um deles apresenta um objetivo final diferenciado. Na mitose, são geradas duas células-filhas, cada uma delas com um número diploide (2n) de cromossomos idêntico ao da célula parental, ao passo que, ao longo da meiose, o número de cromossomos é reduzido à metade da célula-mãe diploide para suas células-filhas haploides (n). Sendo assim, a mitose é amplamente utilizada pelas células eucarióticas somáticas para garantir o crescimento dos tecidos e a reposição de células danificadas ou mortas. Por outro lado, a meiose é essencial para a formação de nossos gametas haploides tomando-se por base a linhagem germinativa, sendo a reprodução sexuada a grande responsável por restaurar o número de cromossomos diploide da espécie a cada nova geração. Neste capítulo, abordaremos como estes dois mecanismos de divisão celular acontecem e como a perturbação destes mecanismos pode levar a doenças clinicamente importantes.

■ Estrutura cromossômica

Cromatina
Complexo formado pela interação entre DNA, RNA, proteínas histonas e não histonas

Antes de descrevermos os detalhes inerentes a ambos os processos de divisão celular, vamos detalhar a estrutura dos cromossomos humanos e de demais componentes essenciais para que esses mecanismos aconteçam de modo preciso.

Os cromossomos nucleares constituem arranjos lineares de genes que se associam a moléculas de RNA e a proteínas histonas e não histonas em diferentes níveis até a formação da rede difusa de filamentos que compõem a cromatina. Ao todo, existem 24 tipos diferentes de

Centrômero
Constrição primária de um cromossomo, que o separa em dois braços. É a posição na qual o cromossomo se liga ao fuso durante a divisão celular, por meio dos cinetócoros

Telômero
Região formada por sequências repetitivas em tandem (TTAGGG)n, localizadas nas extremidades dos cromossomos. Tem o papel de proteger os cromossomos de rearranjos

Origem de replicação
Sítios específicos na molécula de DNA, nos quais proteínas se ligam para dar início à síntese de novas moléculas de DNA

Cromossomo acêntrico
Cromossomo desprovido de centrômero, que é instável durante as divisões celulares, por não poderem se ligar ao fuso

Mitose
Divisão celular realizada pelas células somáticas, na qual uma célula-mãe diplóide dá origem a duas células-filhas diplóides

Meiose
Divisão celular realizada durante a formação dos gametas, na qual uma célula-mãe diploide dá origem a quatro células-filhas haploides, em duas etapas subsequentes (meiose reducional e meiose equacional)

cromossomos humanos (22 cromossomos comuns a ambos os sexos somados aos cromossomos X e Y) que variam em tamanho de 50 a 250 megabases (milhões de bases). Como ilustrado na Figura 4.1, para que um cromossomo seja viável, ele precisa apresentar essencialmente três estruturas: um centrômero, dois telômeros e, no mínimo, uma origem de replicação.

Centrômero ou constrição primária consiste em uma região de intensa espiralização, essencial para a movimentação dinâmica dos cromossomos durante o processo de divisão celular. Cromossomos sem centrômero (acêntricos) são perdidos durante a divisão celular, podendo causar grandes prejuízos às células-filhas.

No início da mitose e da meiose, um arranjo complexo de microtúbulos é formado, e é pela ligação dos centrômeros a esse fuso que os cromossomos podem se movimentar na célula em divisão. Os microtúbulos representam estruturas citoesqueléticas ricas em tubulina que se irradiam de uma pequena organela citoplasmática localizada nos polos da célula, o centrossomo. Essa estrutura representa o centro de organização dos microtúbulos, e é responsável pela coordenação do movimento dos cromossomos. Cada célula apresenta um único centrossomo, formado por um par de centríolos. Durante o processo de intérfase, o centrossomo é duplicado, de modo que cada célula-filha possa receber o próprio centrossomo após a divisão celular. Na divisão celular, os centrossomos-filhos movem-se para polos opostos da célula e novos microtúbulos se desenvolvem a partir de cada centrossomo. Com isso, o fuso pelo qual os cromossomos se movimentarão é estabelecido. Evidências demonstram a participação de proteínas motoras na movimentação dos cromossomos, as quais utilizam energia proveniente da hidrólise de ATP.

Os telômeros (do grego *telos*, final) representam as extremidades dos cromossomos eucarióticos, caracterizadas por uma grande concentração de sequências de DNA repetitivo não codificante, em geral compostas por repetições da sequência TTAGGG que se estendem de 3 kb a 20 kb. Além de serem responsáveis por garantir a estabilidade das extremidades dos cromossomos durante o ciclo celular, minimizando a chance de rearranjos estruturais, os telômeros parecem estar associados à sinalização necessária para o pareamento correto dos cromossomos durante a divisão celular.

O tamanho dos telômeros é mantido por enzimas telomerases, transcriptases reversas ribonucleoproteicas capazes de restaurar o número de repetições presentes nas extremidades cromossômicas. No entanto, a maioria das células somáticas humanas tem quantidade limitada dessa enzima, e a cada vez que a célula se divide ocorre um encurtamento natural das regiões teloméricas, até o ponto em que a célula perde a capacidade de divisão e entra em apoptose. Por esse motivo, o encurtamento dos telômeros tem relação direta com o envelhecimento e

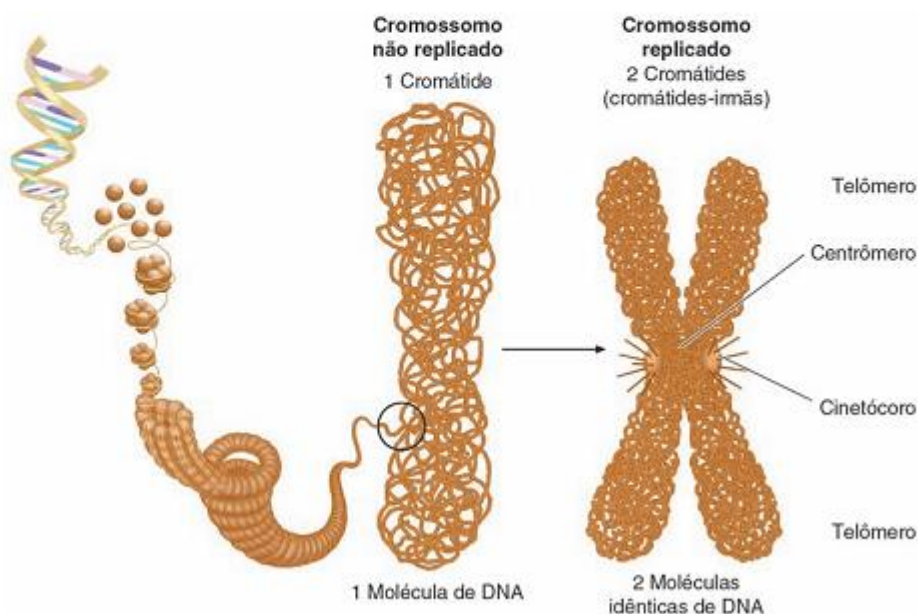



Figura 4.1 Detalhes das principais estruturas observadas ao longo de um cromossomo eucariótico nuclear.

com vários tipos de câncer, já que favorece a ocorrência de rearranjos envolvendo as extremidades dos cromossomos.

Além dos centrômeros e telômeros, é essencial ainda que o cromossomo apresente no mínimo uma origem de replicação, região na qual a duplicação do DNA é iniciada. Antes de se dividir, a célula-mãe precisa duplicar seu conteúdo genético para que as células-filhas resultantes da divisão celular recebam o mesmo conteúdo de informações. A alternância desse processo contínuo formado pelas etapas da intérfase (G1/G0, S e G2) e pela divisão mitótica da célula em si, caracteriza o ciclo celular. Devemos ressaltar, no entanto, que além das fases que compõem o ciclo celular diferirem quanto às suas funcionalidades, a configuração de nossos cromossomos também é distinta nas diferentes etapas, como pode ser observado na Figura 4.2. Na fase G1, quando há produção intensa de RNA e proteínas e a célula se prepara para a replicação de seu material genético, os cromossomos são unifilamentares. A descompactação da cromatina é crescente até alcançar o relaxamento máximo na fase S, quando ocorre a replicação do DNA e a duplicação dos centrossomos. Esse relaxamento é crucial para a ligação dos elementos que compõem a maquinaria replicacional à cromatina, possibilitando a duplicação adequada do conteúdo genético. Na fase S, portanto, os cromossomos não podem ser individualizados por estarem muito finos e distendidos, corando-se pouco.

Centrossomo
Região na qual um par de centríolos organiza a formação dos microtúbulos do fuso durante a divisão celular

A replicação do DNA não é simultânea para todos os cromossomos. Nesse aspecto, o cromossomo X inativo na mulher (Capítulo 2), o qual exibe uma configuração mais compactada em relação aos demais cromossomos, é o que se replica mais tardiamente, característica esta que é amplamente explorada em metodologias investigatórias de anomalias que envolvem os cromossomos sexuais. Tipos celulares que não estão se dividindo afastam-se do ciclo celular, permanecendo em um estado acíclico chamado de G0. Nessa etapa, é verificada redução drástica na velocidade de produção de proteínas em comparação com as células que estão em plena divisão. Esse mecanismo em que a célula permanece em G0 por tempos diferenciados é um dos responsáveis pela variabilidade quanto à duração do ciclo celular entre os diferentes tecidos de nosso organismo. Alguns tipos celulares são estimulados a reingressar no ciclo celular em decorrência de necessidades metabólicas internas e até mesmo mediante estímulos ambientais. Neste último caso, podemos citar como exemplo a exposição demasiada de nossa epiderme à luz solar, o que pode levar a uma necessidade urgente de renovação do tecido epitelial. Por outro lado, alguns tipos celulares não se dividem, como a maioria das células nervosas e das

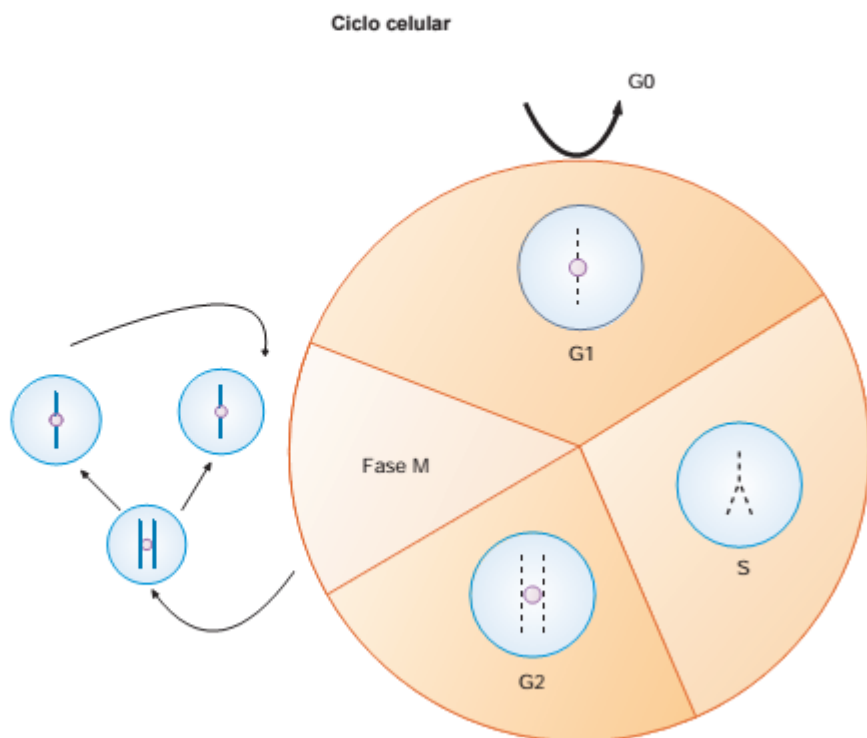



Figura 4.2 Estrutura dos cromossomos ao longo do ciclo celular: intérfase (G1/G0, S e G2) e fase M (mitose). Note que durante a intérfase a estrutura dos cromossomos apresenta-se relaxada, ao passo que, na fase M, ocorre condensação gradativa dos cromossomos.

células do músculo estriado esquelético, as quais permanecerão em G0 permanentemente. Na fase G2, os cromossomos já se apresentam com uma configuração bifilamentar, mas ainda não podem ser reconhecidos de maneira independente, tendo em vista que sua estrutura ainda está relaxada.

Durante a divisão celular, os cromossomos se coram mais facilmente, por estarem mais compactados e absorverem maior quantidade de corantes. No entanto, algumas regiões da cromatina coram-se mais que outras, o que sugere a existência de diferenças na densidade gênica e na atividade transcricional entre as regiões presentes ao longo de um cromossomo. Em geral, as regiões claras (eucromatina) têm menor compactação da cromatina e maior acessibilidade a fatores transcricionais, possibilitando maior atividade dos genes ali presentes. Por esse motivo, nessas regiões é encontrado um grande percentual de genes estruturais. Contrariamente, as regiões mais coradas (heterocromatina) são as que têm cromatina mais compactada, estando, portanto, associadas a regiões de menor atividade gênica.

Eucromatina
Região da cromatina que apresenta estrutura mais relaxada e, por conseguinte, maior concentração de genes em atividade. A configuração descondensada faz que essa região absorva pouca quantidade de corante

Heterocromatina
Região da cromatina que se cora fortemente até mesmo na intérfase, por apresentar um aspecto condensado. A heterocromatina pode ser constitutiva (p. ex.: regiões centroméricas) ou facultativa (um dos cromossomos X na mulher)

Um cromossomo já replicado clássico exibe duas subunidades longitudinais chamadas de cromátides-irmãs, interligadas pelo centrômero. Além disso, com exceção do cromossomo Y, todos os demais cromossomos são encontrados aos pares em nossas células somáticas. Estes, chamados de cromossomos homólogos, compartilham os mesmos *loci*, carregando informações genéticas para as mesmas características hereditárias. Para cada par de cromossomos homólogos, recebemos uma cópia materna e outra de origem paterna. Portanto, cada indivíduo apresenta duas cópias de cada gene, que embora influenciem o mesmo traço, não necessariamente apresentarão a mesma sequência de DNA. Formas alternativas de sequências de DNA para um mesmo gene são chamadas de alelos.

***Locus* (plural, *loci*)**
Posição que um gene ocupa ao longo de um cromossomo. Um único *locus* pode ter diferentes alelos

Mitose

O princípio da mitose está na distribuição igualitária do conteúdo genético presente na célula-mãe diploide para suas duas células-filhas diploides (n=46). Para isso, previamente ao início da mitose, cada um dos cromossomos presentes na célula parental deve estar duplicado, evento que ocorreu durante a fase S da intérfase. Daí, é iniciada a divisão celular mitótica, passando por 5 fases distintas (prófase, metáfase, prometáfase, anáfase e telófase), embora a divisão seja um processo contínuo, conforme demonstrado na Figura 4.3.

Cromossomos homólogos
Par de cromossomos com diferentes origens parentais (um é materno e o outro paterno) que se alinham durante a meiose I e são capazes de trocar material genético entre eles. Os cromossomos sexuais X e Y, ao contrário dos autossomos, apresentam homologia apenas em suas extremidades

A prófase é a fase mais longa da mitose, durante a qual uma condensação crescente dos cromossomos acontece para protegê-los de quebras e rearranjos durante o processo de divisão celular. Nesse estágio, algumas organelas intracelulares se fragmentam (p. ex., complexo de Golgi e retículo endoplasmático) e o nucléolo desaparece. Para que os cromossomos não fiquem dispersos no citoplasma, possibilitando a ocorrência de perdas, os centrossomos migram gradualmente para os polos opostos da célula, dando início à formação do fuso mitótico, ao qual os cromossomos se ligarão. No final da prófase, um par de cinetócoros se forma em cada centrômero, cada um deles ligado a uma cromátide-irmã. Esses elementos representam estruturas proteicas que se associam aos centrômeros e são essenciais para a ligação dos cromossomos ao fuso mitótico.

Alelo
Uma das sequências de DNA possíveis que um gene pode apresentar

A prometáfase é iniciada quando a membrana nuclear se rompe e, com isso, os cromossomos dispersos dentro da célula podem se ligar ao fuso mitótico por meio de seus cinetócoros. Em seguida, os cromossomos começam a se deslocar em direção à porção central entre os dois polos, movimento denominado congressão.

Cinetócoro
Estrutura presente no centrômero, formada pela reunião de proteínas, que serve como ponto de ligação dos cromossomos às fibras do fuso

Congressão
Deslocamento dos cromossomos em direção à porção central entre os dois polos da célula durante a divisão celular

Na metáfase, os cromossomos alcançam o grau máximo de compactação, se tornando mais facilmente corados e identificáveis ao microscópio óptico. Nessa fase, todos os 46 cromossomos estão dispostos na região mediana da célula entre os dois polos (placa equatorial ou placa metafásica). Três tipos de microtúbulos estabilizam os cromossomos na placa metafásica: os microtúbulos polares, cinetocóricos e astrais. Os microtúbulos polares se formam durante a prófase e mantêm os dois polos da célula ligados durante o processo de divisão celular, enquanto os microtúbulos cinetocóricos se desenvolvem no início da metáfase (chamada por muitos autores de prometáfase), sendo os responsáveis pela ligação entre os cinetócoros dos cromossomos metafásicos dispostos na placa equatorial e os polos da célula. Os microtúbulos astrais, por sua vez, são estruturas mais curtas que se irradiam de cada polo da célula.

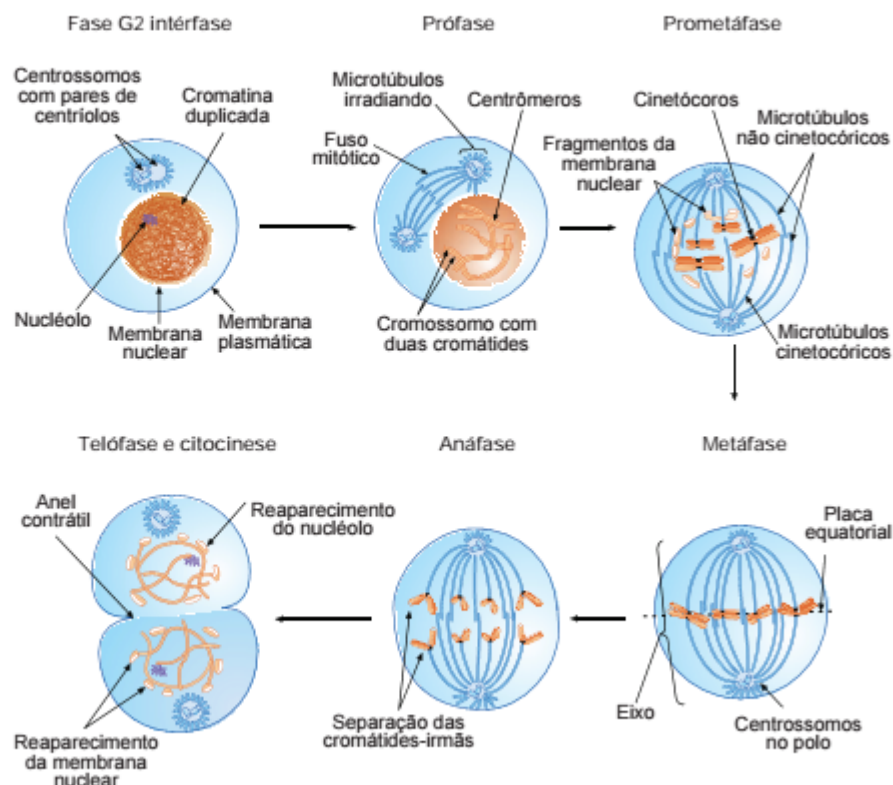



Figura 4.3 Detalhe das diferentes fases que compõem a mitose.

Na anáfase, o centrômero que une as duas cromátides-irmãs se rompe longitudinalmente, fazendo com que as cromátides-irmãs de cada cromossomo sejam levadas para polos opostos da célula. A quebra do centrômero se deve à ação da enzima separase, capaz de dissociar a proteína coesina que mantinha as cromátides unidas desde a duplicação na intérfase. A separação das cromátides-irmãs para polos opostos é chamada de disjunção. Além da quebra proteolítica, o movimento de separação das cromátides-irmãs envolve a participação efetiva dos diferentes tipos de microtúbulos. Nesse sentido, à medida que as cromátides-irmãs se separam, os microtúbulos cinetocóricos começam a se encurtar por despolimerização da tubulina, contraindo o fuso e puxando as cromátides-irmãs para polos opostos. De modo contrário a esse movimento, uma força de repulsão atua nos polos da célula para afastá-los, e os microtúbulos polares são alongados por adição de tubulina, sofrendo deslizamento. O distanciamento entre os polos da célula também tem participação efetiva dos microtúbulos astrais.

Disjunção
Separação das cromátides-irmãs para polos opostos da célula durante a divisão celular

No estágio final da divisão mitótica (telófase), os cromossomos que agora contêm apenas uma das cromátides-irmãs alcançam por completo os polos das células e começam a se descondensar. As fibras do fuso se desintegram, o nucléolo reaparece e todas as demais organelas intracelulares perdidas no início da divisão celular são restauradas. Além disso, cada conjunto de cromossomos em cada polo da célula passa a ser circundado por uma nova membrana nuclear. Assim, a membrana celular da célula-mãe é estrangulada, por meio de um processo denominado citocinese, dando origem a duas novas células-filhas geneticamente idênticas a ela, cada uma delas com uma cromátide-irmã de cada um dos cromossomos originalmente presentes na célula-mãe. A citocinese ocorre pelo acúmulo de actina e miosina ao longo da mitose na linha equatorial da célula, formando um anel contrátil. Ao se contrair no final da telófase, esse anel cria um sulco de clivagem, possibilitando a separação final das células-filhas.

Citocinese
Estrangulamento da membrana celular da célula-mãe na telófase, por meio da formação de um anel contrátil, que torna possível a separação final das células-filhas

Após a citocinese, as células-filhas entram novamente em intérfase, a fim de restaurar a quantidade diploide de material genético presente na célula-mãe. Devemos lembrar, no entanto, que embora o conteúdo genético das células-filhas seja idêntico, curiosamente o quantitativo das estruturas intracelulares pode ser diferente. O desenvolvimento de um organismo multicelular envolve a diferenciação de diversos tipos celulares com base em uma única célula. Um dos mecanismos que possibilita essa diversidade é a capacidade que algumas células têm de se submeter a uma divisão celular assimétrica durante a qual elas segregam determinantes proteicos e RNA nas duas células-filhas de maneira desigual.

Pontos de checagem na mitose

Os pontos de checagem (*check points*) atuam como ferramenta de controle que monitora a execução dos eventos durante o ciclo celular e impede o início de eventos subsequentes até que o processo anterior seja executado com sucesso ou, caso isso não ocorra, a célula entre em apoptose, garantindo a estabilidade genética. A passagem de uma fase para outra é controlada por fatores proteicos evolutivamente conservados que atuam nos pontos de checagem, dentre os quais destacam-se as ciclinas e as quinases dependentes de ciclinas. Em meio aos diferentes pontos de checagem que atuam no ciclo celular, alguns estão diretamente relacionados com a mitose. O primeiro deles ocorre ainda em G2, fase na qual é verificado se a replicação do DNA ocorreu corretamente e a célula pode prosseguir para a fase mitótica. A permissão para o início da mitose é traduzida bioquimicamente pela ativação do complexo promotor de mitose (*mitosis promoting factor* – MPF).

Na mitose propriamente dita, um segundo ponto de checagem não somente monitora a formação do fuso, mas também controla a ligação de suas fibras aos cinetócoros associados aos centrômeros, garantindo que todos os cromossomos sejam capturados pelo fuso. Caso algo esteja errado, um sinal inibitório é propagado e a mitose é suspensa, impedindo a segregação cromossômica desigual para as células-filhas. No final da metáfase, ocorre ainda a ativação do complexo promotor de anáfase (*anaphase promoting complex* – APC), o qual é capaz de controlar a separação das cromátides-irmãs e a migração para polos opostos da célula.

Meiose

Nossos gametas haploides (n=23) são gerados de células diploides (n=46) que compõem a linhagem germinativa. Sendo assim, o princípio da meiose (da palavra grega que significa diminuição) envolve a redução do número de cromossomos à metade em nossos gametas, visando à manutenção do número de cromossomos da espécie. Sem a meiose, o número de cromossomos de nossa espécie duplicaria a cada geração, ultrapassando a capacidade de aporte metabólico da célula.

Antes de ingressar na meiose, as células diploides da linhagem germinativa se multiplicam por meio de sucessivas mitoses. Assim, cada célula resultante da replicação interfásica subsequente terá 46 cromossomos, cada um com duas cromátides-irmãs. Após a multiplicação das células progenitoras dos gametas, a meiose é, então, iniciada e envolve dois processos distintos de divisão celular, sem que haja replicação do conteúdo genético entre eles. Na primeira etapa, ocorre divisão do número de cromossomos pela metade, de modo que a célula-mãe diploide presente na linhagem germinativa (46 cromossomos) gere duas células-filhas portando somente 23 cromossomos cada, ou seja, um cromossomo de cada par presente na célula-mãe. Por essa razão, essa primeira divisão celular da meiose recebe a denominação de divisão reducional. Por outro lado, na segunda divisão da meiose, o número de cromossomos da célula gerada da primeira divisão meiótica (n=23) é mantido para as suas células-filhas (n=23), recebendo o nome de divisão equacional. Nessa etapa, cada cromossomo das células-filhas exibirá uma das cromátides presentes na célula parental. À medida que um gameta feminino e um masculino se unem durante a fertilização, o número diploide (n=46) é restabelecido e o zigoto resultante contendo 46 cromossomos, cada um deles com uma cromátide apenas, sofre sucessivos ciclos celulares para formar o embrião. Assim, a progressão dos episódios que levam à formação dos gametas consiste em: intérfase → divisão reducional (meiose I) → divisão equacional (meiose II), como observado na Figura 4.4.

Divisão reducional
Primeira divisão meiótica que reduz à metade o número de cromossomos da célula-mãe diploide para as células-filhas haploides

Divisão equacional
Segunda divisão meiótica que mantém o número de cromossomos da célula-mãe para as células-filhas. Cada cromossomo das células-filhas exibirá uma das cromátides presentes na célula parental

Meiose I

A primeira divisão meiótica pode ser tradicionalmente representada por quatro fases principais: prófase, metáfase, anáfase e telófase. No entanto, os eventos que acontecem na prófase da meiose I ou, simplesmente, prófase I têm papel de destaque por serem os grandes responsáveis por dar origem à diversidade genética essencial para a sobrevivência de nossa

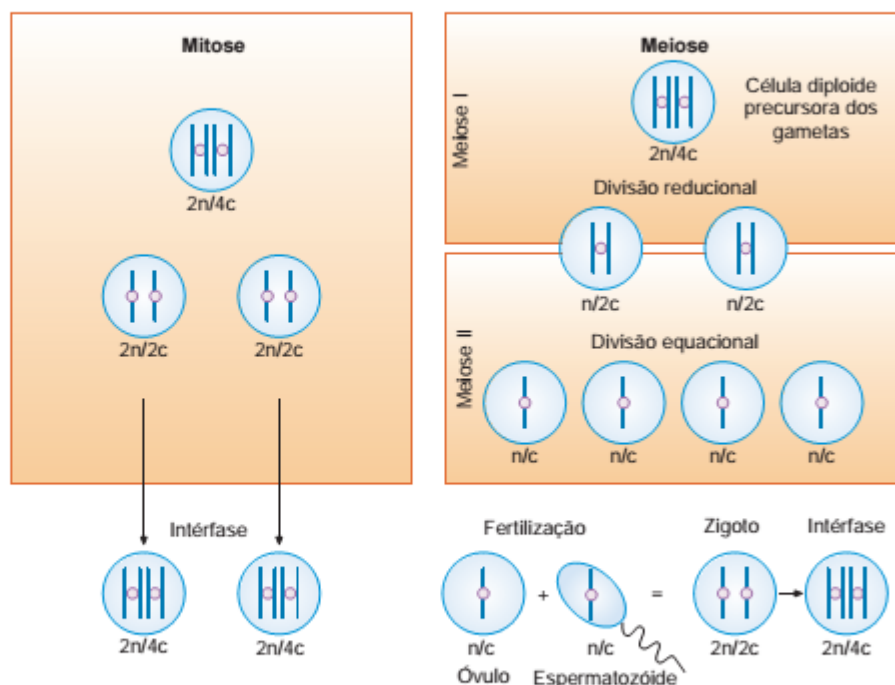



Figura 4.4 As duas divisões meióticas (meiose reducional ou meiose I e meiose equacional ou meiose II) comparadas à mitose. Tanto na meiose quanto na mitose, a célula original é diploide (2n) e tem 4 moléculas de DNA para cada par de cromossomos (4c).

espécie. Em virtude da complexidade e da importância dos eventos que ocorrem na prófase I, esta é subdividida em 5 estágios: leptóteno, zigóteno, paquíteno, diplóteno e diacinese, conforme ilustrado na Figura 4.5.

Durante o leptóteno (do grego, filamentos finos), os 46 cromossomos acabaram de passar por uma replicação na intérfase e, por isso, apresentam uma configuração muito distendida e fina. A espiralização dos cromossomos é crescente dessa primeira etapa da prófase até a diacinese.

No zigóteno (do grego, filamentos pareados), os cromossomos homólogos se alinham lado a lado para formar um bivalente, o qual é composto pela interação de quatro cromátides-irmãs. A esse processo de pareamento que alinha sequências de DNA de cromossomos homólogos dá-se o nome de sinapse e ocorre em vários pontos ao longo do comprimento dos cromossomos homólogos. Como os cromossomos encontram seus homólogos para parearem ainda é uma questão pouco entendida e várias hipóteses foram postuladas. Alguns pesquisadores sugerem que os cromossomos homólogos tendem a ocupar uma mesma posição ainda no núcleo interfásico, o que facilitaria o processo. Outros acreditam que, no final do leptóteno, os telômeros dos cromossomos estariam ligados a pontos adjacentes na membrana nuclear, ocasionando uma configuração polarizada dos cromossomos, o que poderia ter papel relevante na busca pela homologia. De maneira geral, parece que o reconhecimento de um cromossomo por seu homólogo envolve identificadores moleculares específicos para cada um dos cromossomos, localizados ao longo de sua extensão. Uma hipótese seria a de que sequências de DNA repetitivo não codificante representariam pelo menos parte desses identificadores cromossomo-específicos.

Sinapse
Alinhamento entre cromossomos homólogos durante a prófase da primeira divisão meiótica

A sinapse entre homólogos envolve a formação de um arranjo complexo altamente conservado composto por RNA e proteínas, denominado complexo sinaptonêmico. Esse complexo é o responsável pela precisão do processo, por manter os cromossomos fortemente associados durante as próximas etapas. O complexo sinaptonêmico constitui-se de um componente central, basicamente proteico, cercado por componentes laterais dos cromossomos homólogos. Por causa de diferenças em sua composição, durante a formação dos gametas de um indivíduo do sexo masculino, os cromossomos X e Y se associam por regiões restritas de homologia localizadas em suas extremidades (regiões pseudoautossômicas), as quais serão abordadas posteriormente.

Complexo sinaptonêmico
Arranjo complexo altamente conservado, composto por RNA e proteínas, que se forma entre os cromossomos homólogos, durante a prófase da primeira divisão meiótica

Embora o pareamento dos cromossomos homólogos ocorra no zigóteno, é na próxima etapa, o paquíteno (do grego, "filamentos espessos"), que acontecerá um dos eventos de maior impor-

Prófase I
Condensação contínua dos cromossomos

Leptóteno
Cromossomos duplicados tornam-se visíveis

Zigóteno
Pareamento dos cromossomos homólogos

Paquíteno I
Crossing over entre os cromossomos homólogos

Diplóteno I
Repulsão dos cromossomos homólogos e visualização dos quiasmas

Diacinese I
Terminalização dos quiasmas e desmonte do complexo sinaptonêmico

Metáfase I
Cromossomos altamente compactados se alinham na placa equatorial

Anáfase I
Separação dos cromossomos homólogos para polos opostos (disjunção)

Telófase I
Cromossomos homólogos alcançam os polos; restauração da membrana nuclear e nucléolo

Citocinese
Estrangulamento da membrana celular e separação das células-filhas

Prófase II
Condensação dos cromossomos e migração para a placa equatorial

Metáfase II
Cromossomos compactados alinhados na placa equatorial

Anáfase II
Divisão longitudinal dos centrômeros e separação das cromátides-irmãs para polos opostos

Telófase II
Membrana nuclear é refeita e nucléolo é reconstituído

Citocinese
Estrangulamento da membrana celular e formação de 4 células-filhas haploides

Nucléolo

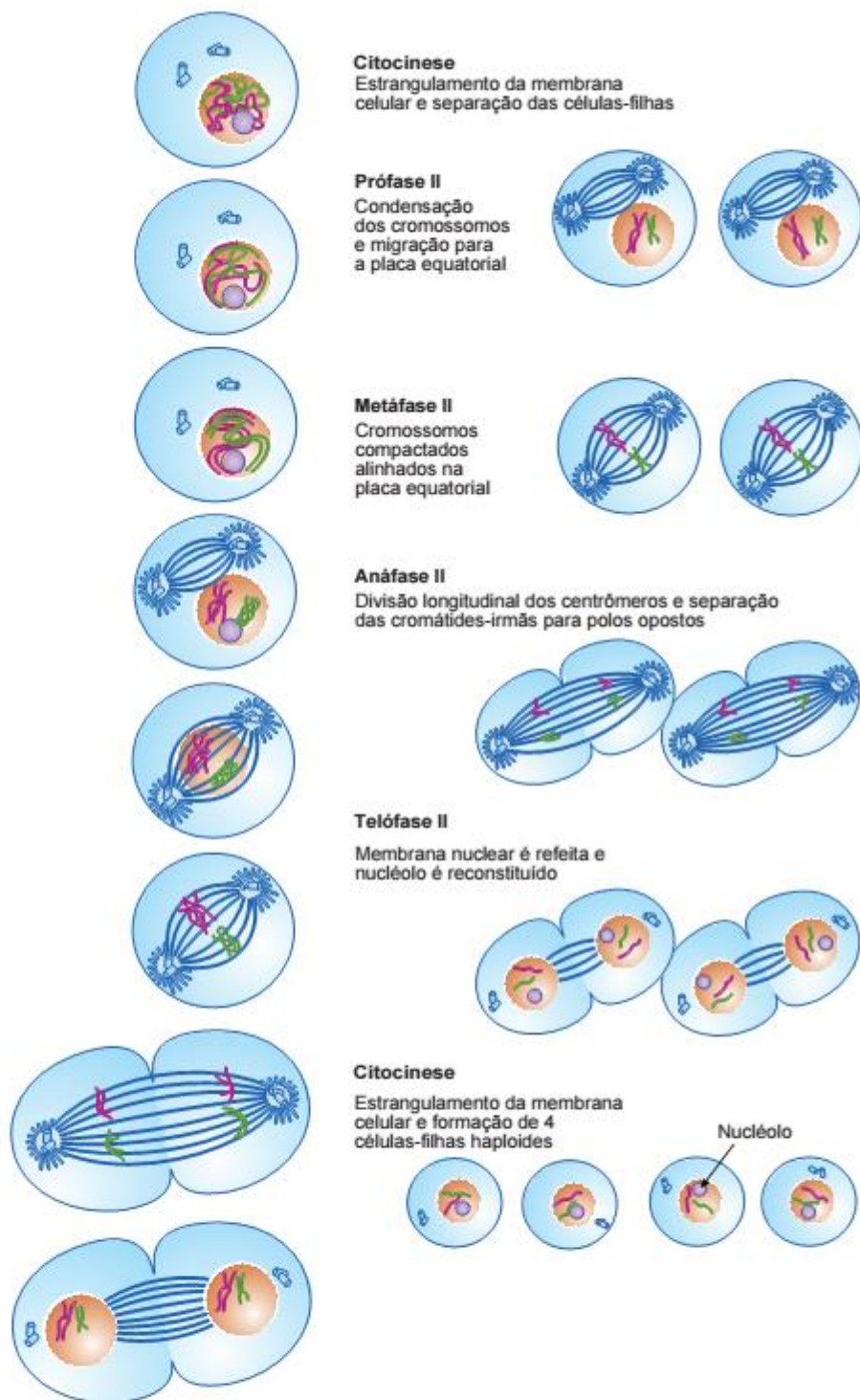



Figura 4.5 Detalhes das diferentes fases que compõem a meiose reducional e a meiose equacional. Note que não há replicação de material genético entre essas duas fases.

tância biológica: o *crossing over*. Nessa etapa, os cromossomos apresentam configuração bem mais condensada e a aparência dos bivalentes se torna mais definida ao microscópio. Pode-se, inclusive, observar a tétrade formada pelo pareamento das quatro cromátides presentes nos cromossomos homólogos, sendo duas cromátides-irmãs do cromossomo de origem paterna e duas cromátides-irmãs do cromossomo de origem materna.

Crossing over
Permuta recíproca de segmentos gênicos entre cromátides não irmãs de cromossomos homólogos

O *crossing over* consiste na recombinação recíproca ou permuta gênica entre cromátides não irmãs de cromossomos homólogos, ou seja, a troca física de material genético entre cromátides homólogas de diferentes origens parentais. Apesar de o *crossing over* acontecer em diferentes pontos ao longo dos cromossomos homólogos, em geral apenas duas cromátides trocam material genético entre si em um determinado ponto. Para que o processo ocorra, as duas cromátides envolvidas devem estar precisamente alinhadas para garantir que nenhuma sequência gênica, mesmo que pequena, seja adicionada ou perdida durante o processo, alterando o funcionamento dos genes envolvidos.

Inicialmente, cada uma das duas cromátides envolvidas no *crossing over* é quebrada em regiões específicas e os segmentos resultantes são trocados entre elas, formando cromossomos recombinantes com novas configurações de sequências gênicas. Por esse motivo, podemos dizer que o *crossing over* é um dos mais importantes mecanismos responsáveis por dar origem à diversidade genética na progênie. É por meio dele que novas combinações gênicas podem ser inseridas em nossos gametas, fornecendo um estoque de variações gênicas à disposição do processo evolutivo. Desse modo, a recombinação garante flexibilidade evolutiva a nosso genoma de realizar um amplo embaralhamento de nossos genes a cada geração, possibilitando a eliminação de variantes desfavoráveis e a inserção de novas combinações de sequências favoráveis. Sem o *crossing over*, essa última tarefa caberia apenas ao processo mutacional, que é um evento raro.

Surpreendentemente, evidências citológicas apontam que o *crossing over* também pode ocorrer na mitose. O *crossing over* acontece na mitose quando segmentos de cromossomos homólogos se pareiam de maneira acidental em células somáticas. Embora seja raro, esse evento parece constituir maneira importante de possibilitar que mutações recessivas associadas ao câncer se expressem.

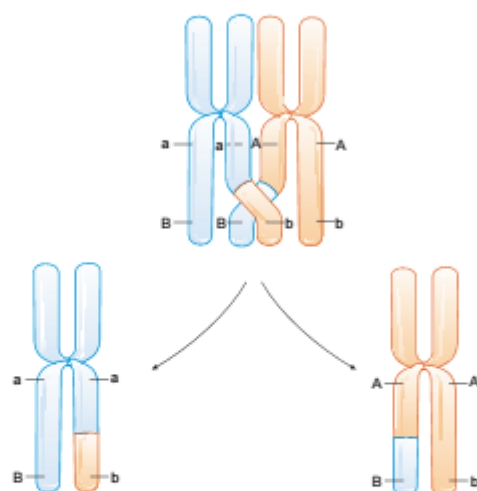
Vários modelos foram desenvolvidos para explicar como o *crossing over* meiótico ocorre, mas a maioria deles é oriunda de dois modelos principais demonstrados na Figura 4.6: a hipótese da quebra unifilamentar em ambas as moléculas, desenvolvido por Holiday em 1964, e a hipótese da quebra bifilamentar, criada por Szostak, Stahl e colaboradores em 1983.

No modelo de Holiday, a clivagem de segmentos correspondentes das duas cromátides envolvidas inicialmente ocorre por meio de uma endonuclease, que corta filamentos únicos da dupla hélice de DNA de cada cromátide-irmã. Em seguida, os filamentos únicos são desenrolados de seus filamentos originais complementares pela ação de DNA-helicases. Então, as extremidades livres dos filamentos únicos simultaneamente se deslocam para o cromossomo homólogo oposto e passam a fazer par com os filamentos complementares intactos presentes nesse cromossomo. Por último, os filamentos clivados já deslocados são ligados a seus novos locais de destino por meio de uma DNA-ligase.

Já no modelo de quebra bifilamentar, a endonuclease cliva inicialmente filamentos duplos em uma das moléculas de DNA participantes e uma lacuna adjacente ao sítio de clivagem é gerada por ação de exonucleases. Essas enzimas removem nucleotídios de cada uma das fitas a partir do ponto de clivagem, gerando terminações 3′ livres. Uma das extremidades 3′ livres recombina-se com um filamento homólogo a ela, presente na segunda molécula de DNA. Nesse momento, ocorre um pareamento de três fitas, chamado de alça D, e a extremidade 3′ livre situa-se entre as duas fitas originais. A alça D é estendida até que todo o espaço produzido na molécula intacta seja preenchido por mecanismos de síntese de reparo do DNA. O deslocamento da alça D faz com que o filamento deslocado hibridize com sequências complementares no cromossomo homólogo e a lacuna de uma das fitas desse cromossomo pode ser restaurada. Em seguida, a lacuna remanescente no outro filamento também é preenchida por complementaridade envolvendo síntese de reparo.

Haplótipo
Conjunto de genes ou variantes alélicas localizadas em um mesmo cromossomo que são transmitidos juntos à próxima geração. Os haplótipos são vastamente utilizados para seguir genealogias e populações, além de serem importantes no mapeamento cromossômico

Ligação genética

Quanto mais próximos dois genes estiverem um do outro, menor a possibilidade de quebra e recombinação entre eles, o que faz com que sejam transmitidos ligados como um único bloco (haplótipo) para o mesmo gameta. Sendo assim, os haplótipos podem ser utilizados como ferramentas para seguir marcadores ligados a genes relacionados com doenças ao longo das

A. Recombinação iniciada por quebras unifilamentares (Holiday, 1964)

1 Cromossomos homólogos

1
2
Cromossomo paterno
3
4
Cromossomo materno

2 Corte unifilamentar em pares homólogos

1
2
3
4

3 Troca entre filamentos simples

1
2
3
4

4 Migração da ramificação

1
2
3
4

5 Corte entre outros segmentos

1
2
3
4

6 Permuta entre os homólogos

1
2
3
4

7 Moléculas recombinantes recíprocas

3
4
1
2

DNA heterodúplex

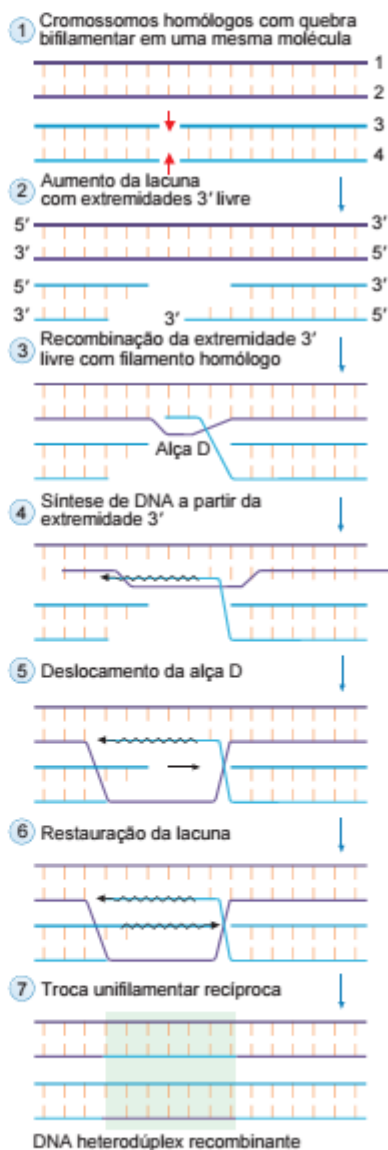



Figura 4.6 Modelos de quebras uni (A) e bifilamentar (B) durante o *crossing over*.

Ligação genética
Transmissão conjunta de dois *loci* para um mesmo gameta durante a meiose, em virtude da grande proximidade entre eles e a ausência de recombinação

Genótipos recombinantes
Novas combinações genotípicas distintas dos genótipos parentais

Loci sintênicos
Loci gênicos localizados em um mesmo cromossomo, independentemente do fato de estarem ligados ou não

Centimorgan
Parâmetro utilizado para medir a distância genética entre dois *loci*. Dois *loci* estarão a 1 cM um do outro se a frequência de recombinação entre eles for de 1%

genealogias ou para seguir populações ao longo do processo evolutivo. Há, então, a ligação genética. Trata-se da transmissão conjunta de dois *loci* para um mesmo gameta durante a meiose, em virtude da grande proximidade entre eles e ausência de recombinação.

A existência de *crossing over* entre dois genes pode ser deduzida pela proporção de indivíduos na progênie que resultam de configurações gaméticas recombinantes (genótipos recombinantes), ou seja, configurações distintas daquelas observadas nos genótipos parentais. Dessa maneira, genes que estão localizados no mesmo cromossomo (sintênicos, que significa "mesmo fio") podem estar fisicamente ligados, impedindo a formação de gametas recombinantes. Por outro lado, para que dois *loci* segreguem de forma independente, eles precisam estar localizados em cromossomos diferentes ou em regiões mais distantes, situadas em um mesmo cromossomo.

Para se calcular a distância genética entre dois *loci* sintênicos, é utilizada a *frequência de recombinação* entre eles, a qual é representada em unidades de *crossing over* ou centimorgans (cM), em homenagem ao cientista Thomas Hunt Morgan, o primeiro a observar um *crossing over* em Drosophila. Dois *loci* em um mesmo cromossomo estão distantes 5 cM um do outro se a chance de eles serem separados um do outro pelo *crossing over* em uma única geração for de 5%. Contudo, devemos ter em mente que a frequência de recombinação entre dois *loci*, por maior que seja a distância física, nunca excederá 50%. Isso acontece porque distância física não é o mesmo que distância genética, pois à medida que a distância física entre dois *loci* aumenta não é evidenciado um aumento proporcional da frequência de recombinação. Nesse sentido, algumas regiões cromossômicas são mais propensas a sofrerem *crossing over* que outras. Regiões adjacentes ao centrômero e próximas às extremidades cromossômicas são menos suscetíveis à permuta gênica. Embora a relação entre distância física e genética não seja constante ao longo de todo o genoma, em regiões de eucromatina, no entanto, em média 1 cM equivale a 1 Mb de DNA.

Para ilustrarmos o conceito de ligação acima descrito, vamos considerar dois *loci* gênicos, denominados *A* e *B*, que apresentam os alelos *A* e *a* e *B* e *b* nos cromossomos parentais (materno e paterno), respectivamente (Figura 4.7). Quando os *loci A* e *B* estão situados em cromossomos não homólogos diferentes, eles segregarão de maneira independente na meiose, gerando 50% de gametas com a mesma configuração cromossômica presente nas células parentais (*AB* ou *ab*) e 50% de gametas com conjuntos cromossômicos com configurações não parentais (*Ab*

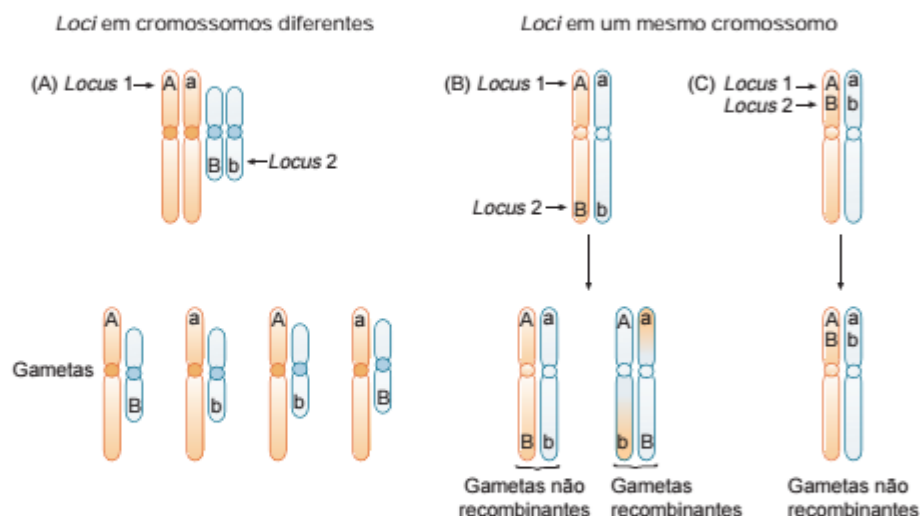



Figura 4.7 Possibilidade de recombinação entre dois *loci* (A e B) de acordo com a localização cromossômica durante a formação dos gametas. (A) Quando os dois *loci* estão localizados em cromossomos diferentes (não homólogos), a segregação de seus alelos (A/a e B/b) durante a meiose ocorre de maneira independente; *loci* localizados em um mesmo cromossomo podem recombinar ou não, dependendo da distância entre eles; (B) *loci* distantes um do outro no mesmo cromossomo sofrem ao menos uma recombinação entre eles, gerando gametas parentais e gametas recombinantes; (C) *loci* localizados muito próximos não sofrem *crossing over* entre eles, originando apenas gametas não recombinantes.

ou *aB*), o que equivale a uma frequência de recombinação de 0,5. Quando os *loci A* e *B* estão fisicamente distantes ao longo de um cromossomo, pelo menos uma permuta gênica acontecerá no segmento entre eles, resultando na presença de genótipos não recombinantes (*A*/*B* e *a*/*b*) e recombinantes (*A*/*b* e *a*/*B*) nos gametas. Caso os dois *loci* estejam muito próximos um do outro, eles serão transmitidos sempre como um único segmento sem recombinação entre eles, gerando apenas genótipos não recombinantes A/B e a/b nos gametas.

Mapeamento genético
Diagrama representativo das posições relativas dos genes ao longo de um cromossomo, com base na frequência de recombinação entre eles, a qual é medida em centimorgan (cM)

Mapeamento físico
Representação gráfica das distâncias (em pares de bases) entre genes e de suas posições relativas em um cromossomo

Duplo heterozigoto
Indivíduo que apresenta heterozigose em cada 1 de 2 *loci* distintos

Lod score
Estimativa estatística para aferir se dois *loci* estão provavelmente próximos um do outro em um mesmo cromossomo e são herdados juntos

A correlação entre a frequência de recombinação e a distância entre dois genes pode ser usada para a construção de um mapa genético, no qual é possível se ordenar a distribuição dos genes ao longo dos cromossomos.

Nesse aspecto, o mapeamento genético se distingue do mapeamento físico de um cromossomo, o qual é elaborado de acordo com a distância medida em pares de bases. No site do National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/mapview), pode ser acessado um banco de mapas cromossômicos para diferentes organismos.

De acordo com o exposto, para definirmos se há ligação entre dois *loci*, o ideal é avaliarmos os indivíduos duplo heterozigotos ao longo de uma genealogia. Quando consideramos a ligação entre dois genes próximos, um indivíduo duplo heterozigoto pode ter os alelos arranjados de duas maneiras nos cromossomos. O modo como os alelos estão dispostos nesses indivíduos heterozigotos é conhecido como fase de ligação. Como na Figura 4.8, por convenção, na configuração *cis* ou de *acoplamento*, os alelos dominantes *A* e *B* se situam em um mesmo cromossomo, ao passo que os alelos recessivos *a* e *b* se localizam no cromossomo homólogo correspondente. Por outro lado, na configuração *trans* ou de *repulsão*, o alelo dominante *A* e o alelo recessivo *b* se situam em um mesmo cromossomo, ao passo que o alelo recessivo *a* e o alelo dominante *B* se localizam no cromossomo homólogo correspondente.

Um dos cálculos estatísticos frequentemente utilizados para confirmar ou afastar a ligação entre dois *loci* pela determinação da frequência de recombinação é o *Lod score* (*log-of-odds ratio*). Esse parâmetro baseia-se na razão logarítmica entre a probabilidade de ligação e a probabilidade de não ligação entre dois genes, de modo que um valor de Lod maior que 3.00 aponta para ligação, ao passo que valores de Lod inferiores a -2.00 indicam ausência de ligação. A utilização desse método estatístico para estudos de ligação tem grande utilidade para a identificação de marcadores genéticos ligados a genes relacionados com doenças. Esses marcadores podem ser representados por variações interindividuais polimórficas, as quais incluem sequências repetitivas *in tandem* (microssatélites) ou polimorfismos de nucleotídio único, localizados próximo a um gene associado a doença humana, podendo ser utilizados no diagnóstico molecular. Para o mapeamento do gene associado a uma doença mendeliana, analisam-se marcadores polimórficos distribuídos pelo genoma com uma distância de 10 cM a 20 cM entre eles em um grande número de famílias de afetados até se identificar um que apresente alelos segregando com a doença.

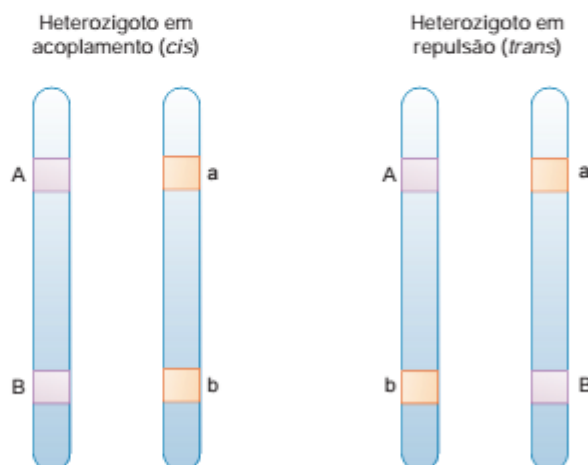



Figura 4.8 Fases possíveis para os alelos A/a e B/b em um indivíduo duplo heterozigoto.

Quiasmas
Evidências citológicas de que o *crossing over* aconteceu durante a meiose I. Podem ser visualizados na fase de diplóteno

Após a etapa de *paquíteno*, na qual os cromossomos realizam o *crossing over*, a célula ingressa na próxima etapa da meiose I, o *diplóteno* (do grego, dois filamentos). Nessa fase, os cromossomos homólogos começam a se afastar um do outro, mas permanecem unidos naqueles pontos onde a recombinação aconteceu. Nessas regiões, os cromossomos exibem uma configuração em forma de X, denominada *quiasma* (da letra X em grego, *chi*), os quais representam as evidências citológicas de que o *crossing over* ocorreu em um dado ponto das cromátides envolvidas.

Na fase de *diacinese* (*dia* = através e *kinesis* = movimento), ocorre a terminalização dos quiasmas, evento caracterizado pelo deslocamento dos quiasmas para as extremidades dos cromossomos e o complexo sinaptonêmico é finalmente desfeito. Além dessas características marcantes, na diacinese, os cromossomos apresentam maior nível de compactação, o nucléolo desaparece e a membrana nuclear é desfeita.

Na *metáfase I*, os cromossomos alcançam o grau máximo de condensação e encontram-se distribuídos na placa equatorial, ligados ao fuso acromático por seus cinetócoros. De modo diferente da que acontece na mitose, a disposição dos cromossomos na placa equatorial na meiose I é distinta, com os homólogos presos ao fuso pelos centrômeros e suas extremidades voltadas para a zona equatorial da célula. Na *anáfase I*, ocorre separação dos cromossomos homólogos, que seguem para polos opostos da célula mediante o encurtamento das fibras do fuso. Esse evento de disjunção dos cromossomos homólogos, assim como o *crossing over*, representa grande fonte de variabilidade para os nossos gametas, e isso se deve ao fato de os cromossomos de origem paterna e materna se alinharem aleatoriamente na placa equatorial durante a metáfase I, de modo que, quando os cromossomos homólogos migram para polos opostos, cada gameta formado contém uma mistura de cromossomos de diferentes origens parentais.

Como temos 23 pares de cromossomos homólogos, existe um total de 2^{23} combinações aleatórias possíveis, ou seja, mais de 8 milhões de possibilidades de células-filhas cromossomicamente distintas. Já na *telófase* I, os cromossomos homólogos separados alcançam os polos da célula, os fusos se desintegram e a membrana nuclear e o nucléolo são restaurados. Por fim, a *citocinese* ocorre visando à separação definitiva das duas células-filhas formadas.

Meiose II

Intercinese
Período de relaxamento da estrutura cromossômica entre a primeira e a segunda divisão meiótica, na qual não ocorre replicação do DNA

Entre a meiose reducional e a equacional não ocorre duplicação do DNA, mas as células passam por um curto período de relaxamento de seus cromossomos, denominado intercinese. A meiose II é, então, iniciada, acontecendo de forma muito semelhante à mitose. A única exceção a essa regra é o fato de, na mitose, a célula-mãe ser diploide (46 cromossomos, cada qual com duas cromátides cada), ao passo que, na meiose II, a célula na qual a divisão se reinicia é haploide (23 cromossomos, cada qual com duas cromátides cada).

Na *prófase II*, ocorre uma condensação gradual dos cromossomos, a membrana nuclear desaparece, assim como o nucléolo. Na etapa seguinte (*metáfase II*), os cromossomos, já compactados ao extremo, alinham-se na placa equatorial. Na *anáfase II*, ocorre divisão longitudinal dos centrômeros e as cromátides-irmãs migram para os polos celulares opostos. Esse processo meiótico de disjunção aleatória das cromátides-irmãs, somado ao *crossing over* e à distribuição aleatória dos cromossomos na anáfase I, também está associado à variabilidade gamética, tendo em vista que as cromátides-irmãs, após a recombinação, já apresentam estruturas gênicas distintas, não sendo mais propriamente "irmãs".

Em seguida à disjunção, as cromátides-irmãs alcançam os polos opostos da célula e a membrana nuclear é refeita na telófase II; subsequentemente, ocorre a citocinese. Sendo assim, após dois processos subsequentes de divisão celular na meiose, uma única célula-mãe origina quatro células-filhas com 23 cromossomos, cada um com uma cromátide apenas.

O número diploide somente é restaurado após a união dos gametas feminino e masculino, originando o zigoto. Este, por sua vez, passa por um processo sucessivo de ciclos celulares até originar o embrião como um todo.

Gametogênese

Dentre os diferentes tipos celulares que nosso organismo apresenta, um grupo de células é direcionado ainda no período embrionário para formar a *linhagem germinativa*, que, posteriormente, originará os nossos *gametas*. O processo contínuo de divisões celulares que envolvem a formação de gametas funcionais é chamado de *gametogênese* (*espermatogênese* para os gametas masculinos e *ovocitogênese* para os gametas femininos) e inclui duas fases principais: a *proliferação* e a *maturação*. Na primeira delas, as *células germinativas primordiais* se dividem por subsequentes divisões mitóticas, visando essencialmente ao aumento substancial do número de células precursoras dos gametas. Após a proliferação, essas células aumentam de volume e preparam-se, então, para a etapa de maturação, na qual as células precursoras se diferenciarão por meiose até a formação final dos gametas (Figura 4.9).

Espermatogênese

A formação dos espermatozoides dura de 64 a 74 dias e acontece nas gônadas masculinas, os *testículos*, em numerosos túbulos finos e espiralizados, chamados de *túbulos seminíferos*. A partir da puberdade, por estímulo hormonal, esses túbulos amadurecem e as células sexuais primárias que os revestem (*espermatogônias*) se proliferam por *mitose*. O enorme número de espermatozoides liberados por um homem a cada ejaculação (cerca de 200 milhões) e produzidos durante toda sua vida (aproximadamente 10^{12}) ocasionam a necessidade de várias centenas de mitoses sucessivas. É válido lembrar que essas espermatogônias estão em diferentes estágios de diferenciação. Sendo assim, em dado momento, um subgrupo de espermatogônias começa a aumentar em tamanho e se diferencia em *espermatócitos primários*. Cada um deles entra, então, em meiose I e dá origem a dois *espermatócitos secundários*, que, por sua vez, rapidamente sofrerão meiose II, originando cada um duas células haploides chamadas de *espermátides*. Até a formação das espermátides, os gametas em desenvolvimento estão interligados por um citoplasma comum durante o processo de divisão celular. Essa característica é fundamental para a eficiência da espermatogênese, pois como a principal característica do gameta masculino é ser pequeno para ter mobilidade, o uso comum do citoplasma fornece o aporte metabólico necessário para a diferenciação final das espermátides em *espermatozoides*. Cada espermátide se diferencia em um espermatozoide por *espermiogênese*. Nesse processo de diferenciação, os espermatozoides podem ser subdivididos em cabeça, pescoço e cauda. Na cabeça, que constitui o maior volume do gameta, podemos encontrar o núcleo com os cromossomos e uma região anterior chamada de *acrossomo*, na qual existem enzimas (hialuronidase, espermiolisina), capazes de digerir a parede do gameta feminino. O acrossomo tem grandes quantidades de complexo de Golgi. No pescoço, região intermediária entre a cabeça e a cauda, podemos encontrar as mitocôndrias, essenciais para a produção de energia e mobilidade à cauda do espermatozoide, formada essencialmente por centríolos.

Ovocitogênese

Ao contrário da espermatogênese, que é iniciada na puberdade e ocorre continuamente durante toda a vida, a formação dos gametas femininos tem seu início ainda no desenvolvimento intrauterino e perdura apenas até a menopausa. Em torno do terceiro mês do desenvolvimento embrionário, as *ovogônias* presentes nos ovários começam a se proliferar por sucessivas mitoses, originando os *ovócitos primários*. Em seguida, os ovócitos primários entram em meiose I até o final da prófase I, quando a divisão celular é suspensa (por ocasião do nascimento) até a puberdade, e a meiose I é reiniciada por estímulo hormonal. A partir desse período, a cada mês em geral um ovócito primário individual irá completar a meiose I, formando um *ovócito secundário* e uma célula contendo pouco volume citoplasmático, chamada de *glóbulo ou corpúsculo polar I*.

Essa fase, na qual os gametas ficam estacionados em prófase I até que completem a meiose I, varia para cada um dos ovócitos e é chamada de dictióteno. É nela que o ovócito adquire um

Dictióteno
Período em que um ovócito primário permanece desde o período do nascimento até a sua completa diferenciação (ovulação)

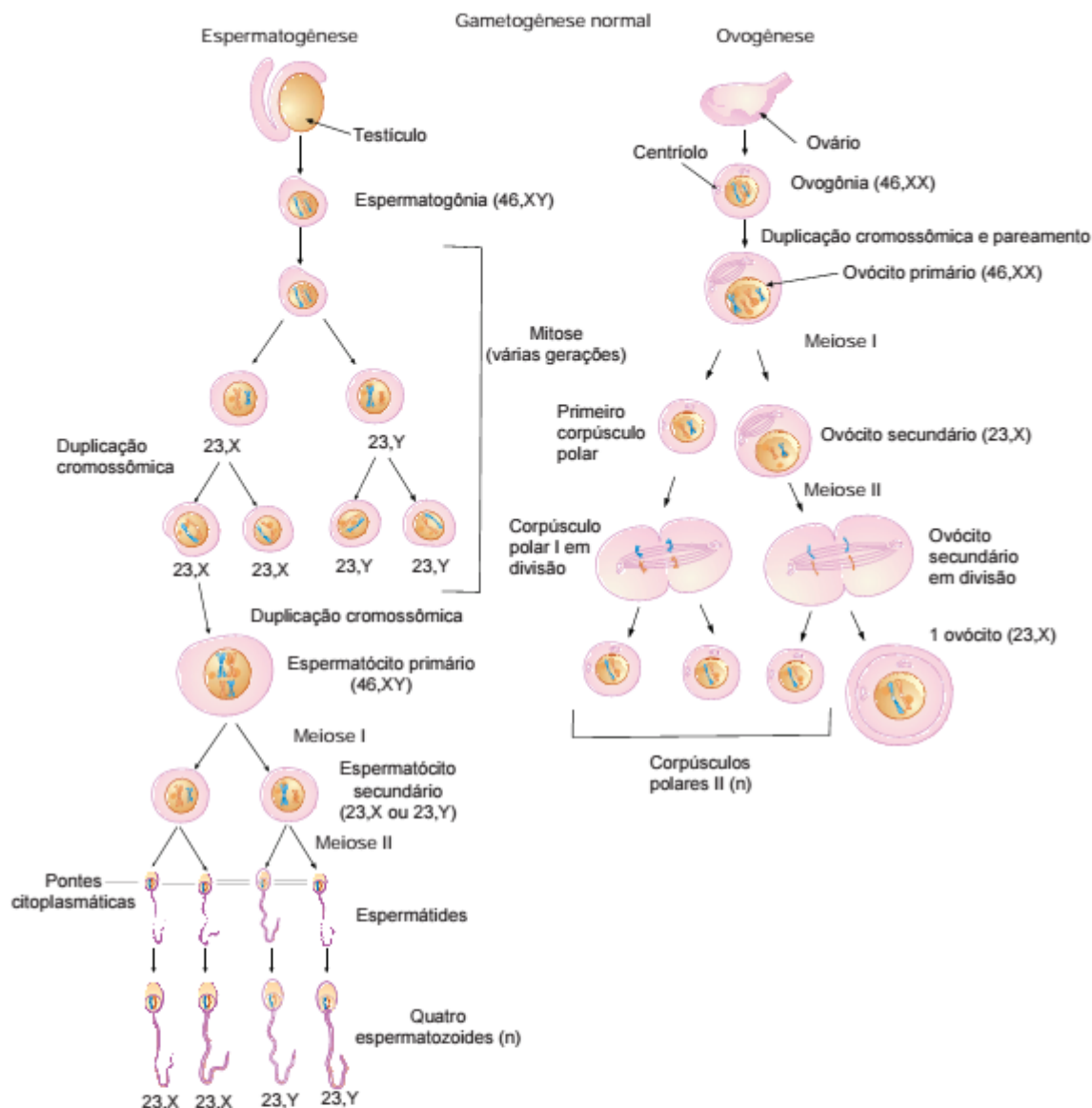



Figura 4.9 Gametogênese masculina e feminina.

revestimento especial que o protege de danos físicos e acumula grande quantidade de nutrientes que serão essenciais para o desenvolvimento embrionário. Além disso, abaixo da membrana celular, desenvolvem-se grânulos corticais que possibilitam que apenas um espermatozoide fecunde cada um dos gametas femininos. À medida que se move ao longo da tuba uterina, o ovócito secundário entra em meiose II até o estágio de metáfase.

Gametas que permanecem em dictióteno por muitas décadas, frequentemente apresentam dificuldades de retornar a meiose II ou podem apresentar falhas nesse processo, levando à formação de embriões com alterações genéticas. Caso a fertilização por um espermatozoide aconteça, o ovócito secundário finaliza a segunda divisão meiótica, o que origina um *segundo corpúsculo polar* e o *óvulo* propriamente dito; do contrário, o ovócito secundário se degenera e à sua liberação mensal é dado o nome erroneamente de ovulação.

Sendo assim, na ovocitogênese, diferentemente da espermatogênese, de uma célula primordial, são gerados um gameta funcional e três corpúsculos polares não funcionais.

Tais corpúsculos representam células muito pequenas que contêm informação genética e capacidade de se dividir, mas não apresentam um número de organelas suficiente para manter

o crescimento de um embrião, visto que são geradas propositalmente por *citocinese assimétrica*. Isso garante ao óvulo a capacidade de nutrir o embrião até que ele possa obter uma fonte própria. A partir do momento em que o espermatozoide e o óvulo se fundem em um núcleo diploide, o zigoto entra em sucessivos ciclos celulares para gerar o embrião como um todo.

Morfologia dos cromossomos

Citogenética
Estudo dos cromossomos

Autossomos
São os 22 cromossomos nucleares comuns a ambos os sexos

Cromossomo metacêntrico
Tipo de cromossomo no qual o centrômero está localizado no centro do cromossomo, gerando braços de tamanhos semelhantes

Cromossomo submetacêntrico
Tipo de cromossomo no qual o centrômero está um pouco deslocado para uma das extremidades, gerando um braço curto e um longo

Cromossomo acrocêntrico
Tipo de cromossomo no qual o centrômero está localizado próximo a uma das extremidades. Os braços curtos desses cromossomos (13, 14, 15, 21 e 22) têm uma segunda região de constrição e transportam genes para RNA ribossômico

Embora o citologista austríaco Walther Flemming tenha sido o primeiro a publicar imagens de cromossomos humanos em 1882, o termo *cromossomo* (corpo colorido) foi introduzido apenas em 1888 pelo citologista alemão W. Waldeyer para evidenciar os filamentos de cromatina que se coravam nas células em divisão. A partir de então, vários pesquisadores da época chegaram a postular que os cromossomos poderiam ser os grandes responsáveis por transportar os determinantes da hereditariedade. Em 1900, com a “redescoberta” das leis de Mendel, Sutton estabeleceu a Teoria Cromossômica da Herança e nomeou de Citogenética a nova ciência que uniria noções prévias da Citologia aos conhecimentos emergentes de Genética. Com a melhoria da resolução dos microscópios óticos no final do século XIX e início do século XX, as atenções dos cientistas se voltaram para a determinação do número correto de cromossomos da espécie humana, assim como para a elucidação das diferenças cromossômicas entre os sexos feminino e masculino. Com esses propósitos, durante várias décadas, o número de cromossomos da espécie humana foi tido como sendo 48. Em 1912, von Miniwarter supôs ainda que os homens teriam 47 cromossomos e que as mulheres teriam 48. Foi finalmente em 1956, com a otimização das técnicas de cultivo celular disponíveis, que Tjio & Levan, ao estudarem células removidas do tecido pulmonar de embriões humanos, elucidaram que o número correto de cromossomos de nossa espécie era 46. Com essa descoberta, consolidou-se a Citogenética e, a partir de então, houve necessidade crescente de unificação da nomenclatura utilizada para a descrição dos cromossomos humanos, visando não somente a comunicação entre diferentes grupos de citogeneticistas ao longo do mundo, mas também a caracterização das alterações cromossômicas que começaram a ser descritas. Essa padronização foi inicialmente obtida com a publicação do *International System for Human Cytogenetic Nomenclature* (ISCN) em 1978 e vem sendo periodicamente atualizada. Tomando-se por base as padronizações estabelecidas ao longo dos anos, os 44 cromossomos humanos encontrados em ambos os sexos foram denominados autossomos e foram numerados aos pares de 1 a 22, enquanto os dois cromossomos sexuais restantes foram chamados de XX nas mulheres e XY nos homens, totalizando os 46 cromossomos que compõem o conjunto diploide de nossa espécie.

Segundo o ISCN, os cromossomos humanos organizam-se sob 3 parâmetros: *tamanho*, *forma* e *padrão de bandas*. Quanto ao tamanho, classificam-se em pares de 1 a 22, em ordem decrescente de tamanho, em 7 grupos de A a G. A única exceção a essa regra é o cromossomo 21, que é ligeiramente menor que o cromossomo 22. Contudo, quando essa classificação foi estabelecida, o cromossomo 21 já era historicamente relacionado com a síndrome de Down e convencionou-se manter a antiga designação. O cromossomo X é de tamanho intermediário e o Y se assemelha em tamanho ao cromossomo 22.

Quanto à *forma*, os cromossomos metafásicos humanos classificam-se em *metacêntricos*, *submetacêntricos* e *acrocêntricos*, como demonstrado na Figura 4.10.

Essa classificação está baseada na posição que o centrômero ocupa ao longo do cromossomo, dividindo-o em dois segmentos, chamados de *braço curto* (*p*, do francês *petit*) e *longo* (*q*, letra em sequência a letra p no alfabeto). Nos cromossomos metacêntricos, o centrômero está localizado na região central, dividindo-o em dois segmentos de tamanho muito semelhante. Nesse caso, embora não haja distinção notável entre o tamanho dos braços, as abreviaturas p e q são usadas como referência aos braços de cima e de baixo, respectivamente. Nos cromossomos submetacêntricos, o centrômero encontra-se deslocado para uma das extremidades do cromossomo, gerando braços claramente de tamanhos diferentes. Por último, nos cromossomos acrocêntricos, o centrômero encontra-se muito afastado para uma das extremidades. Nesse último subgrupo de

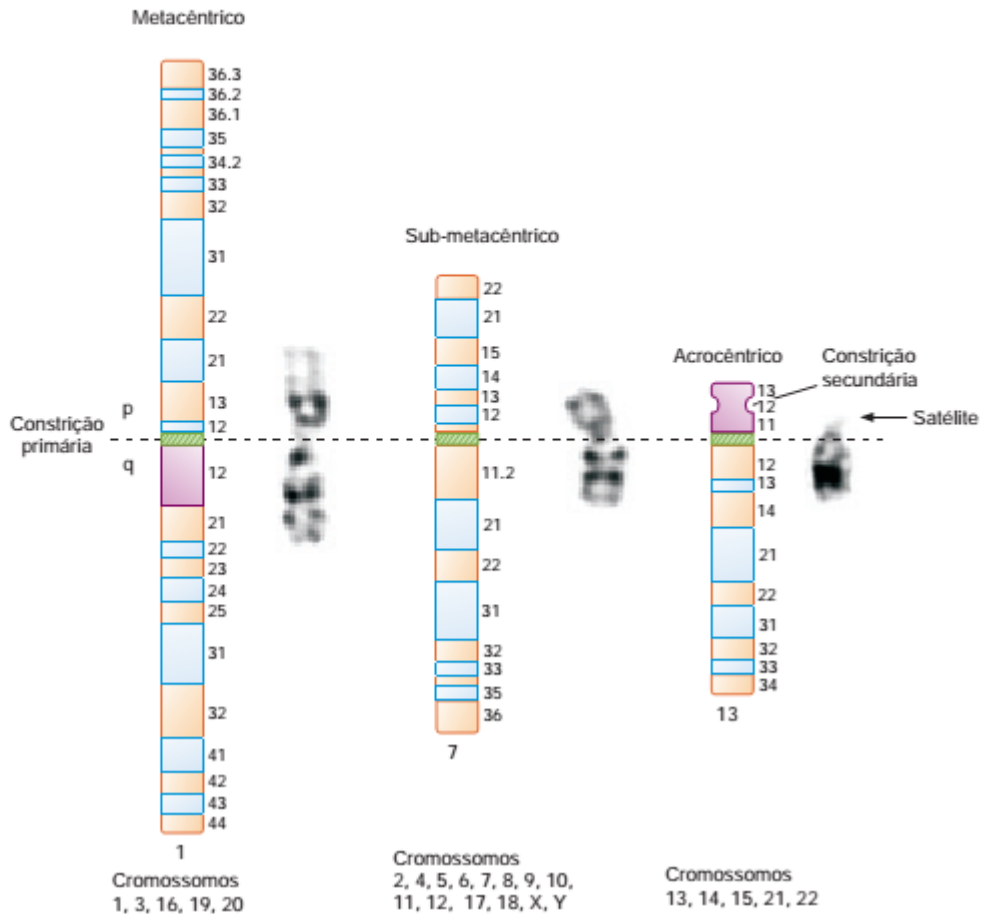



Figura 4.10 Classificação dos cromossomos quanto à posição do centrômero.

cromossomos, além da constrição primária da cromatina representada pelo centrômero, podemos observar uma segunda região de extrema condensação (constrição secundária). As constrições secundárias dos cromossomos acrocêntricos atuam como *regiões organizadoras de nucléolos*. Os nucléolos são organelas arredondadas responsáveis pela produção e processamento de RNA ribossômico e proteínas ribossômicas. No início da divisão celular, os nucléolos que se encontram localizados nas constrições secundárias dos cromossomos acrocêntricos se desprendem e migram para o citoplasma, sendo responsáveis pela organização dos ribossomos, locais onde ocorrerá a produção de proteínas. A parte distal à região onde se encontravam os nucléolos é chamada de satélite. Em algumas espécies, tais como de salamandras e camundongos, mas não na espécie humana, pode ser encontrado ainda um quarto subgrupo de cromossomos, denominado telocêntricos. Nesse caso, o centrômero está localizado exatamente em uma das extremidades e o cromossomo em questão só apresenta um braço.

Com base na utilização de técnicas de coloração adequadas (bandeamento cromossômico), é possível distinguir, ao longo de cada braço cromossomo, uma alternância entre regiões claras e escuras. Isso se deve parcialmente a diferenças de condensação do DNA entre regiões eucromáticas e heterocromáticas. De acordo com esse padrão, as diferentes regiões observadas podem ser numeradas a partir do centrômero, em direção aos telômeros como p1, p2, p3, e assim por diante para o braço curto, e q1, q2, q3 etc. para o braço longo. As regiões, por sua vez, podem ser divididas em bandas, as quais recebem a identificação p11 (um, um e não onze), p12, p13 etc., sendo ainda subdivididas em sub-bandas como p11.11, p11.12, p11.13 etc. Com base nas características de tamanho, da forma e da análise minuciosa desse padrão de bandeamento cromossômico, torna-se possível a distinção de cada um dos cromossomos e a identificação de alterações que os envolvam. A ordenação dos cromossomos de um indivíduo aos pares e em ordem decrescente de tamanho recebe o nome de cariótipo, como pode ser observado na Figura 4.11.

Constrição primária
Região de intensa espiralização presente em todos os cromossomos viáveis. Representa o mesmo que centrômero

Constrição secundária
Segunda região de intensa espiralização, observada nos cromossomos acrocêntricos. Nessas regiões, situam-se os genes que produzem RNA ribossômico

Satélite
Região distal à constrição secundária

Cromossomo telocêntrico
Tipo de cromossomo no qual o centrômero está localizado exatamente em uma das extremidades, de maneira que o cromossomo só apresenta um braço

Bandeamento cromossômico
Grupo de técnicas citogenéticas promovem a coloração diferenciada de cromossomos ou regiões cromossômicas, possibilitando a identificação de alterações cromossômicas

Cariótipo
Constituição cromossômica de um indivíduo, na qual os cromossomos são ordenados aos pares e em ordem decrescente de tamanho

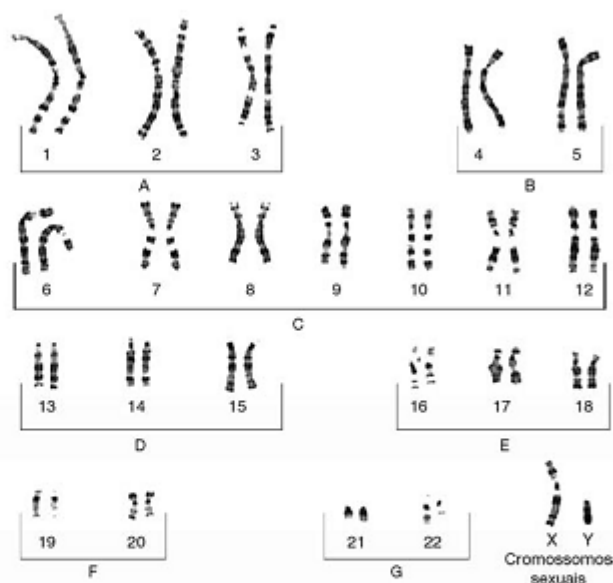



Figura 4.11 Cariótipo normal de um homem (46,XY).

Cariótipo e técnicas de bandeamento cromossômico

Como descrito, os cromossomos são melhor visualizados ao microscópio óptico durante a divisão celular, principalmente na fase de metáfase, que é a etapa na qual eles se encontram mais compactados. No entanto, é mais fácil se estudar os cromossomos humanos em mitose, tendo em vista que a visualização de cromossomos humanos em meiose só é possível pelas amostras de ovário ou testículo. Na meiose feminina, a dificuldade é mais extrema já que grande parte dessa divisão celular acontece ainda no ovário fetal. Por outro lado, a obtenção de cromossomos mitóticos metafásicos pode ser realizada com base em vários tecidos, incluindo fibroblastos da pele, células da medula óssea, linfócitos do sangue periférico, biopsias de tumores sólidos, além de vilosidades coriônicas, células do cordão umbilical e células fetais livres no líquido amniótico utilizadas para o diagnóstico pré-natal. No entanto, os linfócitos ainda representam o tecido de escolha na maioria dos casos, em virtude não só de sua melhor acessibilidade, como também da facilidade quanto ao cultivo *in vitro*.

Após a exclusão de fatores ambientais associados a uma determinada condição, as principais *indicações clínicas* que devem ser consideradas para a análise citogenética são: atraso no desenvolvimento físico e/ou intelectual, dismorfias faciais, malformações congênitas, baixa estatura, genitália ambígua, atraso no desenvolvimento de características sexuais secundárias, histórico de infertilidade ou abortos espontâneos recorrentes, natimortos, dentre outros. Além disso, a *idade materna avançada*, a existência prévia de uma criança com alteração cromossômica e a presença de algum rearranjo cromossômico em um dos genitores também devem ser considerados como indicativos da realização de exame citogenético.

Na análise citogenética envolvendo linfócitos humanos, uma amostra de sangue periférico é coletada e misturada a um meio enriquecido com nutrientes adequados, além de uma substância chamada de *fito-hemaglutinina*, que atua como um indutor mitótico em potencial. A cultura é então incubada por 70 h, ao final das quais se adiciona *colchicina* a elas. Esse elemento se liga especificamente à tubulina dos microtúbulos, interrompendo a divisão celular em metáfase, o que possibilita a análise dos cromossomos ao microscópio ótico após o rompimento da membrana nuclear e uma série sucessiva de lavagens e fixações. Em algumas situações, se faz necessária a interrupção da cultura na fase que antecede a metáfase (*pró-metáfase*), utilizando-se para esse fim algumas substâncias capazes de se intercalar ao DNA como o brometo de etídio. Com isso, os cromossomos passam a exibir uma configuração menos condensada com a sua visualização em maior resolução, o que possibilita a identificação de pequenas alterações ao longo dos cromossomos. No entanto, na coloração densa padrão utilizada para visualização dos cromossomos, não é possível a *individualização* de cada um deles. Para esse fim, as lâminas contendo os cromossomos são submetidas a diferentes tratamentos envolvendo a desnaturação e/ou digestão enzimática dos cromossomos, que após coloração apropriada possibilitam a distinção dos cromossomos. São reconhecidos vários tipos de bandeamentos cromossômicos, dos quais o mais popular é o *bandeamento G*. Esse bandeamento cria, após a digestão proteolítica dos cromossomos com tripsina e coloração com corante Giemsa,

um padrão de bandas alternado em regiões claras e escuras, próprio para cada um dos 24 tipos de cromossomos diferentes de nossa espécie (22 autossomos somados aos cromossomos X e Y). Outros padrões de bandeamento incluem: o *bandeamento R* (reverso), no qual o padrão de bandas obtido com tratamento com calor é o inverso daquele observado no bandeamento G; o *bandeamento C*, no qual apenas a região centromérica dos cromossomos é corada após o tratamento com hidróxido de bário e o *bandeamento NOR*, que utiliza tratamento com nitrato de prata para a visualização das regiões organizadoras de nucléolos, presentes nos cromossomos acrocêntricos.

Nas duas últimas décadas, com o avanço das técnicas de Biologia Molecular, a análise Citogenética adquiriu um novo formato, a Citogenética Molecular, a qual vem tornando a análise dos cromossomos humanos cada vez mais precisa e menos artesanal. Dentre as metodologias desenvolvidas, podemos citar a hibridação *in situ* com fluorescência (FISH), que possibilita a visualização de cromossomos inteiros ou partes específicas de um cromossomo pela utilização de sondas moleculares fluorescentes em preparações de cromossomos metafásicos desnaturados e até mesmo no núcleo interfásico. Essas sondas representam sequências de DNA que têm alta homologia com a região cromossômica de interesse e quando se ligam especificamente a ela, a sua fluorescência pode ser identificada por meio de microscópio apropriado. O aprimoramento da técnica FISH tem possibilitado ainda a identificação simultânea de várias sequências-alvo, por meio da utilização de várias sondas, marcadas com fluoróforos diferentes. Muitas vezes, o número de sondas exige o uso de filtros de imagem e softwares especializados para a análise dos cromossomos. Dois procedimentos, o FISH multiplex (M-FISH ou multiplex FISH) e o cariótipo espectral (SKY, *spectral karyotype*), são utilizados nesse sentido, diferindo apenas no método utilizado para discriminar as sondas diferencialmente marcadas. Essas sondas são capazes de rastrear o genoma humano por inteiro, tendo aplicações que vão desde o mapeamento de novos genes até o diagnóstico de alterações cromossômicas. Recentemente, uma nova técnica, denominada 3-D FISH ou *FISH on 3D-preserved nuclei*, tem combinado a técnica FISH à microscopia 3D e à reconstrução de imagem, possibilitando a análise do arranjo espacial de sequências de DNA no núcleo da célula e o estudo detalhado da estrutura e função celulares. A organização tridimensional dos cromossomos helicoidizados dentro do núcleo parece variar entre os diferentes tipos celulares e reflete uma ordem superior de regulação epigenética, sendo considerada a nova fronteira da genômica (Figura 4.12).

FISH
Hibridação *in situ* com fluorescência

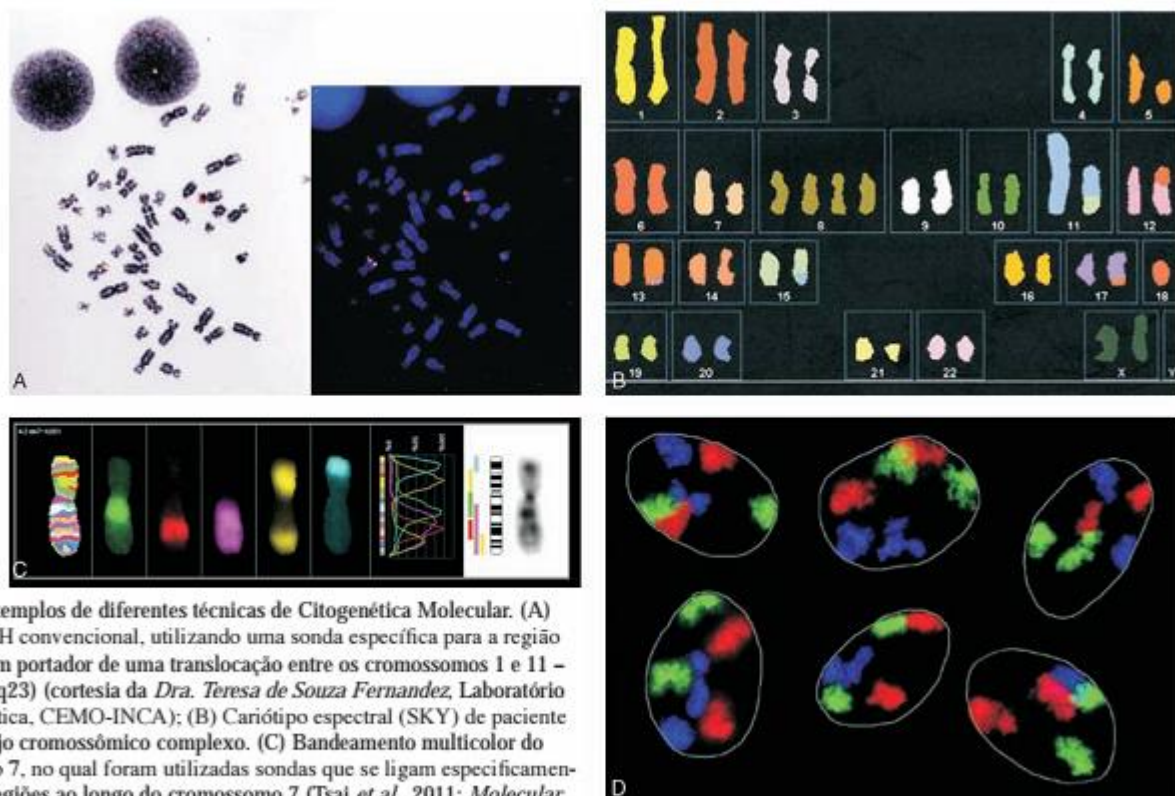



Figura 4.12 Exemplos de diferentes técnicas de Citogenética Molecular. (A) Técnica FISH convencional, utilizando uma sonda específica para a região 11q23 em um portador de uma translocação entre os cromossomos 1 e 11 – t(1;11)(q21;q23) (cortesia da *Dra. Teresa de Souza Fernandez*, Laboratório de Citogenética, CEMO-INCA); (B) Cariótipo espectral (SKY) de paciente com rearranjo cromossômico complexo. (C) Bandeamento multicolor do cromossomo 7, no qual foram utilizadas sondas que se ligam especificamente a várias regiões ao longo do cromossomo 7 (Tsai *et al.*, 2011; *Molecular Cytogenetics* 2011, 4:5). (D) Imagem de 3D-FISH com três cores distintas em núcleos de fibroblastos humanos diploides normais (http://cshprotocols.cshlp.org/cgi/content/full/2007/5/pdb.prot4730/F2).

Alterações cromossômicas

A manutenção da estrutura e morfologia dos cromossomos humanos é fundamental para o correto funcionamento celular. Como os cromossomos consistem em arranjos lineares de genes, qualquer alteração em seu número e/ou estrutura pode alterar a expressão de genes importantes.

As alterações cromossômicas subdividem-se em *anomalias cromossômicas numéricas*, quando há mudanças no número padrão de cromossomos, ou *estruturais*, se há alterações na morfologia de um ou mais cromossomos.

O *International System for Human Cytogenetic Nomenclature* (ISCN) desenvolveu inúmeras abreviaturas que facilitam a descrição dessas alterações envolvendo os cromossomos humanos e podem ser resumidamente vistas na Tabela 4.1.

Tabela 4.1 Algumas das abreviaturas utilizadas nos cariótipos humanos.

Abreviatura	Significado	Exemplos	Especificações
cen	Centrômero	–	–
del	Deleção	46,XY,del(7p)	Homem com deleção no braço curto do cromossomo 7
dup	Duplicação	46,XX,dup(11q)	Mulher com duplicação no braço longo do cromossomo 11
der	Cromossomo derivado	46,XX,der(15)	Mulher com cromossomo resultante de uma translocação, derivada do cromossomo 15
dic	Cromossomo dicêntrico	46,X,dic(Y)	Homem com cromossomo Y contendo dois centrômeros
fra	Sítio frágil	46,Y,fra(X)(q27.3)	Sítio frágil na região q27.3 do cromossomo X, associado à síndrome do X-frágil
i	Isocromossomo	46,X,i(Xq)	Mulher com síndrome de Turner em decorrência de isocromossomo do braço longo do X
ins	Inserção	46,XX,ins(16;6)(p12;p21.2p23)	Mulher com inserção intercromossômica da região 6p21.2 a p23 no cromossomo 16 em 16p12
inv	Inversão	46,XX, inv(3)(p25q21)	Mulher com inversão pericêntrica do cromossomo 3
mar	Cromossomo marcador	47,XY,+mar	Homem com cromossomo extra não identificado
mat	Origem materna	47,XX,+21 mat	Mulher com cromossomo 21 extra de origem materna
p	Braço curto	–	–
pat	Origem paterna	47,XX,+18pat	Mulher com cromossomo 18 extra de origem paterna
q	Braço longo	–	–
r	Cromossomo em anel	46,X,r(X)	Mulher com um dos cromossomos X em anel
rob	Translocação robertsoniana	46,XY, rob(13;21)(q10;q10)	Homem com translocação robertsoniana entre os cromossomos 13 e 21. Os pontos de quebra e troca aconteceram na região q10 de ambos os cromossomos
t	Translocação	46,XY,t(14;21)	Translocação recíproca entre os cromossomos 14 e 21
tel	Telômero	–	–
+	Adição	47,XY,+13	Homem com trissomia do 13 ou síndrome de Patau
–	Perda	45,XX,-22	Mulher com monossomia do 22
/	Mosaicismo	45,X/46,XX	Mulher com dois tipos de linhagens celulares

Anomalias cromossômicas numéricas

Euploidia
Alteração cromossômica que resulta em um número de cromossomos que é múltiplo exato do conjunto haploide

As *alterações cromossômicas numéricas* podem ser classificadas em euploidias e aneuploidias. As *euploidias* são alterações que levam a célula a ter um número de cromossomos que é um múltiplo exato do conjunto haploide, ao passo que as aneuploidias envolvem um ou mais cromossomos específicos, dando origem a células com um número de cromossomos que não é múltiplo do conjunto haploide característico. Os principais tipos de *euploidia* são:

- *Nuliploidia*: são células desprovidas de todo o conjunto de cromossomos. Apesar de um organismo nuliploide não ser viável, algumas células nuliploides podem ser encontradas

em situações fisiológicas normais em nosso organismo. Embora a maioria de nossas células seja diploide, algumas células extremamente diferenciadas, como as plaquetas e hemácias, são anucleadas e desprovidas de cromossomos

- *Haploidia* (n): são células, nas quais o conjunto de cromossomos encontra-se em dose única, como acontece normalmente nos gametas. No entanto, nas células somáticas de nosso organismo diploide, esta condição é anormal e inviável
- *Poliploidia* (3n, 4n etc.): acontece quando o conjunto de cromossomos está representado 3 (*triploidia*, n=69 cromossomos), 4 (*tetraploidia*, n=92 cromossomos) ou mais vezes. Embora as poliploidias sejam comuns em plantas, tendo grande importância em seu processo evolutivo, em humanos essa condição é encontrada em 1% a 3% das gestações humanas, levando a abortos espontâneos em geral até o quarto mês de gestação. Células poliploides podem ser encontradas em células tumorais. As triploidias são mais frequentes que as tetraploidias, sendo ainda observados alguns poucos casos, nos quais as gestações chegam a termo, embora os indivíduos triploides não sobrevivam após esse período.

Como ilustrado na Figura 4.13, a *triploidia* pode ser causada por dois eventos diferentes. No primeiro e mais comum deles (2/3 do total de casos), uma falha na função dos grânulos corticais, localizados abaixo da membrana celular do óvulo, faz com que ele seja fertilizado por dois espermatozoides (dispermia). No segundo mecanismo, um dos gametas (feminino ou masculino) é diploide, em decorrência de um erro meiótico. Pode acontecer inclusive a incorporação de um corpúsculo polar no ovócito em formação. A manifestação fenotípica da triploidia dependerá da origem parental do conjunto extra de cromossomos. Quando o conjunto extra de cromossomos é de origem paterna, o desenvolvimento da placenta é anormal, representando o que se chama de *mola hidatiforme parcial*. Quando o conjunto de cromossomos a mais for de origem materna, o embrião será abortado precocemente durante a gestação.

Dispermia
Fertilização de um ovócito por dois espermatozoides

Na *tetraploidia*, a principal causa é a *endomitose*, que corresponde a duas replicações sucessivas do DNA após a fertilização sem divisão mitótica entre elas (Figura 4.13). A presença de constituições tetraploides sexuais 92,XXXX ou 92,XXYY e a ausência das combinações 92,XXXY ou 92,XYYY comprovam esta etiologia.

Embora a *poliploidia constitucional* seja um evento prejudicial à viabilidade do organismo, algumas das células funcionais de nosso organismo podem ser poliploides sem que isso traga consequências fenotípicas negativas. Por exemplo, os hepatócitos do fígado são submetidos

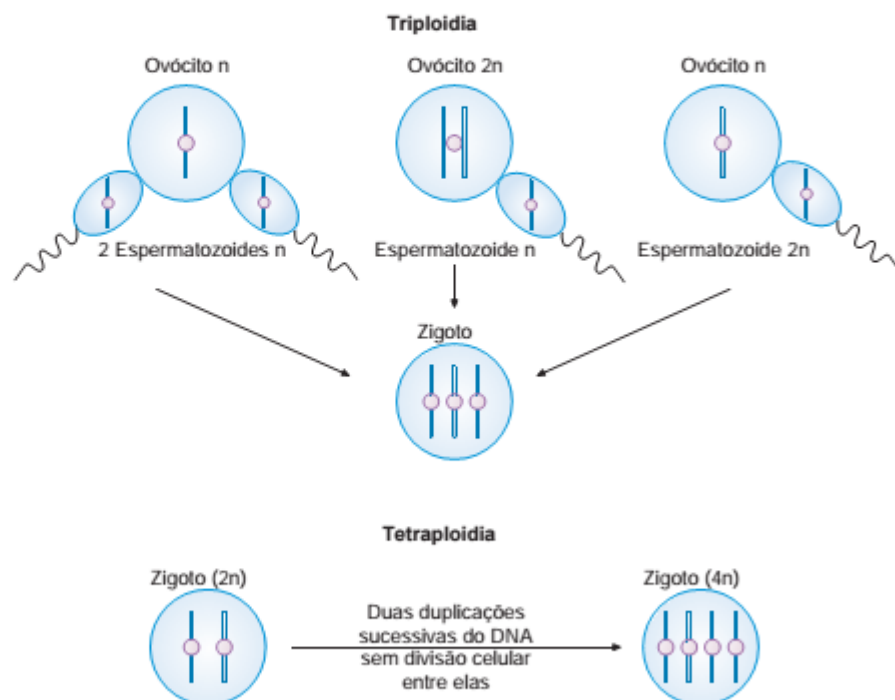



Figura 4.13 Principais causas para o surgimento de triploidia e tetraploidia.

Aneuploidia
Alteração cromossômica que resulta em um número de cromossomos que não é múltiplo exato do conjunto haploide. As formas mais comuns de aneuploidia são as trissomias (adição de um cromossomo extra) e as monossomias (perda de um cromossomo de um par de homólogos)

Monossomia
Alteração cromossômica numérica que surge em consequência da presença de apenas uma cópia de um cromossomo, em vez das duas habituais

Trissomia
Alteração cromossômica numérica que surge em consequência de três cópias de um mesmo cromossomo, em vez das duas previstas

frequentemente à endomitose, em virtude de sua ampla capacidade de regeneração celular. Os megacariócitos que dão origem a milhares de plaquetas nuliploides, o número de cromossomos pode ser de 8, 16 ou 32 vezes o conjunto haploide.

Em contraste com as *euploidias*, as *aneuploidias* são alterações que não envolvem todo o conjunto de cromossomos da célula, mas, sim, cromossomos específicos.

As aneuploidias em humanos ocorrem em pelo menos 5% de todas as gestações e podem contribuir para mais de 1/3 de todos os abortos espontâneos. As principais aneuploidias são:

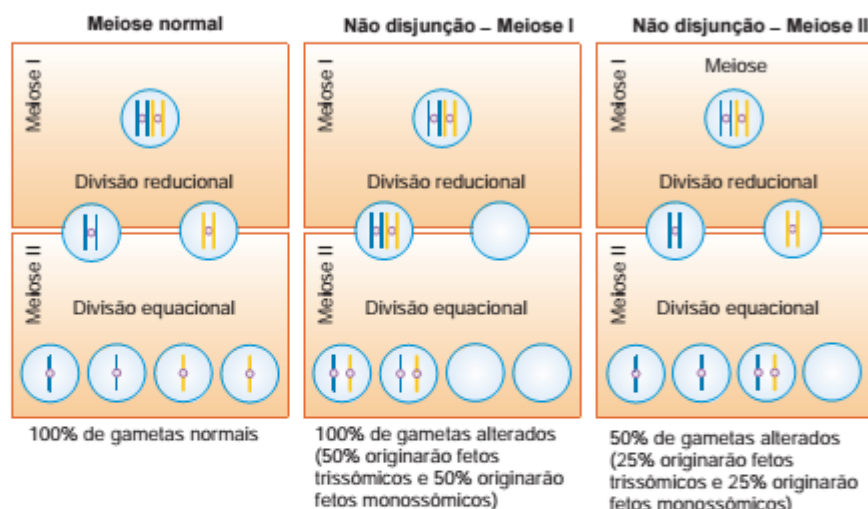
- *Nulissomia* (2n-2): ocorre quando os dois membros de um mesmo par de cromossomos são perdidos. Em geral, essas alterações são inviáveis ainda no período de pré-implantação
- *Monossomia* (2n-1): é uma condição na qual um cromossomo específico está presente em apenas uma cópia. Com exceção da monossomia do cromossomo X em mulheres (45,X ou síndrome de Turner), em geral as monossomias completas dos demais cromossomos são incompatíveis com o desenvolvimento. Em células tumorais, no entanto, monossomias completas podem ser observadas, como é o caso da monossomia do cromossomo 21, identificada em alguns indivíduos com leucemia mielomonocítica aguda e leucemia mielomonocítica crônica. Por outro lado, monossomias parciais que correspondem à perda de parte de um cromossomo específico são viáveis e podem estar relacionadas com características sindrômicas, como é o caso da síndrome da monossomia da região 1p36. Nessa síndrome, a criança apresenta deficiência intelectual de moderada a grave, atraso no crescimento, convulsões, capacidade de fala limitada, hipotonia, perda auditiva e de visão, além de malformações distintas
- *Trissomia* (2n+1): ocorre quando um cromossomo em específico está presente em três cópias em vez de duas, como era de se esperar. Em geral, o aumento da expressão de genes em decorrência de trissomias é menos prejudicial para o organismo como um todo do que a perda gênica resultante das monossomias. Por esse motivo, algumas trissomias são viáveis e apresentam grande importância sob o ponto de vista clínico, causando defeitos congênitos bem reconhecíveis. As principais trissomias conhecidas são as trissomias dos cromossomos 13, 18 e 21, associadas às síndromes de Patau, Edwards e Down, respectivamente, e a trissomia do cromossomo X, ou síndrome do triplo X.

Outras aneuploidias mais raras incluem a *tetrassomia* e a *trissomia dupla*. Na tetrassomia (2n+2), um mesmo cromossomo está presente em quatro cópias. A condição só é viável quando envolve o cromossomo X (48,XXXX), por causa da inativação dos cromossomos X extras. Já na trissomia dupla (2n+1+1), ocorre trissomia de dois cromossomos de pares cromossômicos distintos. Como exemplo, podemos citar a presença de trissomia do cromossomo 21 concomitante à trissomia do cromossomo X ou síndrome do triplo X, originando um indivíduo que apresenta o cariótipo 48,XXX,+21.

Quando as aneuploidias envolvem o aumento do número total de cromossomos, conforme notado nas trissomias, tetrassomias e trissomias duplas, ela é denominada *hiperdiploide*. Quando essas alterações contribuem para uma diminuição do número de cromossomos (nulissomia e monossomia) são chamadas de *hipodiploides*.

As aneuploidias resultam da *não disjunção*, isto é, a não separação dos cromossomos durante a anáfase, podendo acontecer nas anáfases meióticas I ou II durante a formação dos gametas ou na anáfase da mitose após a formação do zigoto.

A não disjunção pode ser primária, quando o indivíduo é cromossomicamente normal e produz gametas anormais ou secundária, quando o indivíduo é cromossomicamente anormal e produz gametas alterados. A não disjunção meiótica dos cromossomos na gametogênese é um dos principais mecanismos que contribuem para o desenvolvimento de um feto cromossomicamente anormal em nossa espécie. Conforme demonstrado na Figura 4.14, caso o erro aconteça na primeira divisão meiótica, o gameta que apresentará o cromossomo extra terá dois cromossomos de um mesmo par em vez de uma cópia apenas. Além disso, os dois membros do par terão origens parentais distintas (um será de origem materna e o outro paterna). Quando a não disjunção acontece na meiose II, o gameta contendo o cromossomo em excesso tem duas

Figura 4.14 Representação da não disjunção para um determinado par de cromossomos durante a primeira e a segunda divisão meiótica. Repare na proporção de gametas alterados e na origem parental diferenciada dos gametas contendo o cromossomo extra.

cópias de um mesmo cromossomo que apresentam a mesma origem parental (ou materna ou paterna). Quando o erro acontece na primeira divisão meiótica, todos os gametas resultantes serão afetados, enquanto em caso de não disjunção na meiose II, apenas 50% deles o serão. Devemos lembrar que os mesmos mecanismos que levam à inserção de uma cópia cromossômica extra nos ovócitos e espermatozoides também podem levar à ausência de cromossomos específicos nesses gametas, possibilitando o surgimento de monossomias após a fecundação por um gameta normal.

Mosaicismo
Presença de duas ou mais linhagens celulares em um mesmo tecido ou em diferentes tecidos do organismo originados de um único zigoto

Quimeirismo
Presença de duas ou mais linhagens celulares em um mesmo tecido ou em diferentes tecidos do organismo originados de pelo menos dois zigotos diferentes

Quando a não disjunção acontece na anáfase mitótica após a formação do zigoto, o indivíduo gerado apresentará mais de uma linhagem celular. Nesse aspecto, dois mecanismos distintos podem ser evidenciados: o mosaicismo por aneuploidia e o quimeirismo, como ser observado na Figura 4.15. No mosaicismo, duas ou mais linhagens celulares são observadas em um mesmo tecido ou em diferentes tecidos do corpo, sendo provenientes de um mesmo zigoto. Quando o erro na formação dos gametas acontece em decorrência de uma não disjunção mitótica, essas linhagens se distinguem entre si quanto ao número de cromossomos (p. ex.: 46,XX/47,XX,+21). Nesse caso, a proporção de células atingidas pelo erro em cada tecido ou entre diferentes tecidos dependerá do momento em que a não disjunção aconteceu durante o

Figura 4.15 Mosaicismo e quimeirismo. Repare que, no mosaicismo (A), as diferentes linhagens celulares foram geradas a partir de um único zigoto, enquanto, no quimeirismo (B), as diferentes constituições celulares foram formadas após a união de dois zigotos.

desenvolvimento embrionário. Além disso, a presença da linhagem normal pode minimizar as consequências fenotípicas da alteração, muitas vezes aumentando até mesmo a viabilidade do indivíduo. Devemos ressaltar que o conceito de mosaicismo também pode ser extrapolado para alterações estruturais e para as euploidias.

Na mitologia grega, o termo quimera representa um ser constituído por segmentos de diferentes animais. Na Genética, o termo quimeirismo é utilizado quando se observa se a presença de duas ou mais linhagens celulares em um mesmo indivíduo, provenientes de pelo menos dois zigotos diferentes. Uma das hipóteses para explicar tal situação está na mistura de células embrionárias provenientes de zigotos distintos, formando um único indivíduo. O *quimeirismo constitucional* (verdadeiro) parece ser um evento raro, sendo mais aparente em caso de fusões envolvendo embriões de sexos opostos, o que causa grandes consequências no desenvolvimento. Em alguns casos, no entanto, o que se observa é um *microquimeirismo*, que envolve o intercâmbio entre células sanguíneas durante a gestação de gêmeos dizigotos. Como consequência, os gêmeos apresentarão células sanguíneas geneticamente de dois tipos. O mesmo evento também já foi relatado durante gestações únicas, nas quais algumas células do embrião passam para a circulação materna e vice-versa, ocorrendo permanência dessas células em ambos os indivíduos, mesmo que por vários anos após o parto. Outro evento que poderia favorecer a ocorrência de quimeirismo é a fertilização *in vitro*. Para aumentar as chances de sucesso, nesse procedimento é comum que mais de um embrião seja implantado na mãe, o que pode aumentar a incidência de fusões embrionárias. O termo quimerirismo também tem sido amplamente empregado nos dias atuais em procedimentos que envolvem clonagem, nos quais o material genético de um indivíduo é introduzido em uma célula tronco embrionária pluripotente. Sob estímulo adequado, essa fusão origina um embrião formado por uma quimera de células do blastocisto e do indivíduo que doou o núcleo.

Causas das alterações cromossômicas numéricas

O principal fator etiológico associado à não disjunção cromossômica tanto de autossomos como de cromossomos sexuais durante a formação dos gametas é a *idade materna avançada*, que pode fazer com que um ovócito em dictióteno apresente dificuldades em finalizar as divisões meióticas após vários anos. No entanto, grande parte das crianças com diferentes aneuploidias nasce de mães com menos de 35 anos. Um dos mecanismos que poderia explicar este evento seria a presença de polimorfismos em genes que participam do *metabolismo do folato*, como o polimorfismo 677C>T no gene da enzima metilenotetra-hidrofolato redutase (*MTHFR*). O folato é uma vitamina do complexo B obtida pela dieta, principalmente com a ingestão de miúdos de animais e com vegetais folhosos verde-escuros, e está envolvido em vias bioquímicas essenciais, como na biossíntese de nucleotídios, no reparo do DNA e nas reações de metilação. A diminuição da biodisponibilidade do folato por meio da sua baixa ingestão e/ou da presença de polimorfismos em genes envolvidos no seu metabolismo poderia levar à *hipometilação do DNA*, inclusive das *regiões centroméricas e pericentroméricas*, expressão gênica anormal, segregação meiótica alterada e acúmulo de metabólitos neurotóxicos. Sendo assim, polimorfismos em genes que codificam enzimas envolvidas no metabolismo do folato têm sido associados a um risco aumentado de não disjunção cromossômica materna em várias populações, principalmente nos casos de síndrome de Down. Todavia, esses polimorfismos podem ter em parte suas consequências metabólicas negativas neutralizadas pela ingestão adequada de folato e micronutrientes a ela relacionados. Assim, se considerarmos que a maioria dos casos de não disjunção ocorre na meiose I materna que acontece no ovário fetal, a suposta prevenção dos eventos de não disjunção em uma gestação atual deveria ter sido previamente realizada pela avó da criança que está sendo gerada, na ocasião da gestação de sua mãe.

Estudos sobre a influência da *idade paterna* sobre a presença de aneuploidias nos espermatozoides são inconsistentes e, em geral, a presença de aneuploidias nos gametas masculinos é atribuída a fatores como a exposição demasiada a pesticidas organofosfatados e tratamento quimioterápico. No entanto, recentemente foi demonstrado que homens com níveis aumentados de folato poderiam apresentar uma frequência reduzida de diferentes tipos de aneuploidias no esperma.

Exemplos de alterações cromossômicas numéricas

A seguir, estão listadas as principais síndromes associadas a alterações cromossômicas numéricas autossômicas. As alterações numéricas relacionadas com os cromossomos sexuais serão abordadas separadamente mais adiante.

Trissomia do cromossomo 21 ou síndrome de Down

A *síndrome de Down* (47,XX,+21 ou 47,XY,+21) é um perfil metabólico complexo causado pela presença de uma cópia extra total ou parcial do *cromossomo 21* no embrião, sendo reconhecida como a trissomia mais comumente observada em humanos e a causa genética mais frequente de deficiência intelectual. A incidência da síndrome é de 1 em 600 nativivos e 1 a cada 150 concepções, representando importante causa de aborto espontâneo, com uma perda fetal estimada de 80% durante a gravidez. Em 95% dos casos, o evento da não disjunção ocorre primariamente durante a meiose I da ovocitogênese materna.

As características clínicas da síndrome incluem: habilidade cognitiva abaixo da média, hipotonia, baixa estatura, pescoço curto, orelhas pequenas com baixa implantação, olhos com fissuras palpebrais oblíquas, pregas epicânticas (prega cutânea que recobre o canto interno do olho), macroglossia (língua grande, protusa e sulcada), encurvamento dos quintos dígitos, aumento da distância entre o primeiro e o segundo artelho, prega única transversal nas palmas das mãos (prega simiesca) e microcefalia. Os portadores da síndrome de Down têm ainda um risco aumentado de apresentar doenças cardíacas congênitas (presente em 1/3 dos pacientes), disfunções tireoidianas, leucemia e malformações do trato gastrintestinal. Além disso, pacientes adultos com síndrome de Down, a partir dos 40 anos, apresentam, frequentemente, alterações cerebrais atróficas, características da *doença de Alzheimer*. Esse perfil demencial pode ser explicado, em parte, pela presença de três cópias do gene *APP*, que está localizado no cromossomo 21 e codifica a proteína precursora de beta-amiloide, peptídio que se acumula no encéfalo de pacientes com doença de Alzheimer. Curiosamente, pacientes com essa doença também parecem exibir maior frequência de células com trissomia do cromossomo 21, em decorrência de erros mitóticos do que o esperado.

O excesso de material genético proveniente do cromossomo 21 em pacientes com síndrome de Down pode surgir de quatro formas distintas:

- trissomia do cromossomo 21 em todas as células do indivíduo por não disjunção meiótica, representando mais de 90% dos casos

Genética em Foco

Expressão gênica global na síndrome de Down

Nos últimos anos, com o sequenciamento do cromossomo 21, o desenvolvimento de modelos animais para a síndrome de Down e a disponibilização de novas ferramentas metodológicas, os mecanismos moleculares envolvidos na síndrome de Down têm sido progressivamente esclarecidos. Nesse sentido, a análise global da expressão gênica tem demonstrado que o pleiotropismo (quadro fenotípico variado) na síndrome de Down poderia ser causado pela superexpressão de genes que não estão localizados ao longo do cromossomo 21, mas, sim, ao longo do genoma. Isso indica que provavelmente na região 21q22.2–22.3 existem genes reguladores, cuja expressão em triplicata na síndrome de Down levaria à desregulação de outros genes distribuídos ao longo dos demais cromossomos. Enquanto o quebra-cabeça não é completamente montado, é consenso que as formas de tratamento de portadores com a síndrome de Down têm evoluído bastante, o que se reflete diretamente no aumento da expectativa de vida desses pacientes, a qual passou de 12 anos em 1949 para mais de 60 anos atualmente.

- trissomia do cromossomo 21 em algumas células do indivíduo por não disjunção mitótica em adição a uma linhagem celular normal (mosaicismo somático) que está presente em 2% dos casos
- translocação de uma terceira cópia do cromossomo 21 em outro cromossomo acrocêntrico (mais frequentemente os cromossomos 14 ou 21). Essa alteração está presente em 4% dos pacientes com a síndrome e será detalhada mais adiante
- trissomia parcial do cromossomo 21, na qual pelo menos a região crítica para a síndrome de Down (21q22.2–22.3) está em triplicata. Esses casos são bem raros, mas de grande importância científica, pois são de ampla utilidade na identificação de regiões específicas do cromossomo 21, responsáveis por uma característica fenotípica em particular da síndrome.

Trissomia do cromossomo 13 ou síndrome de Patau

A *síndrome de Patau* (47,XX,+13 ou 47,XY,+13) é caracterizada pela existência no embrião de três cópias do *cromossomo 13*, causando sério comprometimento do sistema nervoso central durante o desenvolvimento embrionário e morte prematura, em geral, até o primeiro mês de vida. A incidência da síndrome é de 1 a cada 15.000 a 20.000 nascimentos e as características clínicas aparentes são: malformações cardíacas, holoprosencefalia, fenda labial e palatina, defeitos oculares graves (microftalmia, ausência de olhos, coloboma da íris), polidactilia, pés virados para fora e calcanhar proeminente (chamado de pé de cadeira de balanço), defeitos urogenitais, hemangiomas e deficiência intelectual grave nos indivíduos que sobrevivem por um período maior de tempo. Além da trissomia livre do cromossomo 13 e do mosaicismo somático, a etiologia da síndrome também pode incluir, em 20% dos casos, translocações não balanceadas envolvendo o cromossomo 13 extra.

Trissomia do cromossomo 18 ou síndrome de Edwards

A *síndrome de Edwards* (47,XX,+18 ou 47,XY,+18) é causada pela presença de uma cópia extra do *cromossomo 18*, sendo a segunda trissomia autossômica mais frequente, com uma incidência de 1 a cada 7.500 nascidos vivos. As principais manifestações clínicas da síndrome incluem: hipertonia, boca pequena, fissuras faciais acentuadas, micrognatia, mãos cerradas de maneira característica (segundo e quinto dedos sobrepostos aos demais), pés em formato de cadeira de balanço, orelhas de baixa implantação, osso occipital saliente, esterno curto, alterações renais (rim em ferradura) e malformações cardíacas em mais de 50% dos casos. Grande parte dos fetos com trissomia do 18 é abortada espontaneamente e a sobrevida dos portadores da síndrome que chegam a termo é baixa com morte, em geral, na primeira semana de vida. Os indivíduos que sobrevivem apresentam comprometimento intelectual grave. Em alguns casos, a presença de trissomia do cromossomo 18 em forma de mosaicismo somático e de outros cariótipos atípicos pode aumentar a viabilidade da síndrome.

Anomalias cromossômicas estruturais

As alterações que envolvem a estrutura dos cromossomos resultam de uma ou mais rupturas transversais em um ou mais cromossomos, seguida de uma reconstituição cromossômica anormal. Essas alterações podem ocorrer espontaneamente em uma baixa frequência ou serem induzidas pela ação de agentes externos, tais como a exposição à radiação ionizante, uso de drogas mutagênicas e infecções virais (rubéola, citomegalovírus etc.). Assim como as alterações numéricas, as alterações cromossômicas estruturais também podem afetar todas as células do organismo ou apenas uma proporção delas (mosaicismo somático ou germinativo), podendo estar restritas aos gametas.

Os rearranjos estruturais classificam-se em *balanceados*, se a reconstituição cromossômica após quebra transversal não provocar perda ou adição de material genético, ou *não balanceados*, se houver alteração quantitativa do conteúdo genético.

Em geral, é esperado que rearranjos balanceados não originem um fenótipo anormal. No entanto, como uma exceção a esta regra existem aqueles rearranjos que são aparentemente ba-

lanceados ao microscópio ótico (nível citogenético) e que em um nível molecular apresentam-se desbalanceados. Nesse sentido, uma quebra cromossômica em um rearranjo citogeneticamente balanceado pode, por exemplo, afetar a expressão de um gene, seja por meio da interrupção de sua sequência codificadora ou da separação desse gene de seu elemento de controle. Pode acontecer ainda de o gene envolvido na quebra cromossômica ser translocado para um ambiente no qual a configuração da cromatina está diferente daquela presente em sua posição original (mais relaxada ou mais compactada), afetando drasticamente as taxas de transcrição gênica (Capítulo 2). De acordo com o abordado, alterações aparentemente balanceadas que levam a um fenótipo anormal são de grande relevância para a identificação de genes que, quando mutados, estão associados a doenças humanas. Devemos lembrar ainda que os rearranjos balanceados podem resultar em configurações cromossômicas desbalanceadas na gametogênese graças a erros no pareamento dos cromossomos homólogos.

Sendo assim, progenitores portadores de rearranjos balanceados fenotipicamente normais podem originar progênie com constituições cromossômicas desbalanceadas.

Quando uma ruptura transversal ocorre em um cromossomo são formadas duas extremidades reativas, e em seguida três eventos subsequentes podem ocorrer: (a) as extremidades podem se reunir, restaurando a constituição original do cromossomo, sem maiores prejuízos à célula e ao organismo como um todo; (b) as extremidades reativas não se reúnem e com isso ocorre perda do segmento acêntrico (sem centrômero) durante a divisão celular subsequente pelo fato do cromossomo não conseguir se prender ao fuso. Paralelamente, o segmento cromossômico que permaneceu com o centrômero fica com uma deficiência de material cromossômico; (c) o segmento resultante da ruptura transversal se une a outro cromossomo, o que é chamado de translocação.

As alterações estruturais podem estar presentes em apenas *uma* ou *ambas* as *cromátides*. Para que a quebra esteja presente em ambas as cromátides (*alteração estrutural cromossômica*), é necessário que o evento tenha acontecido em uma das cromátides antes da célula se replicar, provavelmente na *fase G1 da intérfase*. Dessa forma, na fase S, a nova molécula que será sintetizada com base na molécula de DNA alterada também apresentará a mesma falha. Quando a alteração estrutural envolve apenas uma cromátide (*alteração estrutural cromatídica*), o erro ocorre após a molécula de DNA já ter sido replicada, em geral na *fase G2 da intérfase*.

Além disso, os *rearranjos estruturais* também classificam-se como *intracromossômicos*, quando afetam apenas um cromossomo, ou *intercromossômicos*, quando mais de um cromossomo está envolvido no rearranjo.

Alterações estruturais intracromossômicas

As principais alterações intracromossômicas conhecidas são as deleções (del), duplicações (dup), inversões (inv), isocromossomos (i), cromossomos em anel (r), cromossomos dicêntricos (dic) e cromossomos marcadores (mar). Como podemos observar na Figura 4.16, as deleções cromossômicas representam quebras transversais que provocam perda de material cromossômico contido no segmento acêntrico. As manifestações clínicas em portadores de deleções dependerão da extensão do material cromossômico contido no segmento perdido, bem como da importância dos genes ali presentes. Em geral, as consequências fenotípicas em casos de deleções resultam da incapacidade de uma única cópia do(s) gene(s) deletado(s) não ser(em) suficiente(s) para manter as funções originalmente desempenhadas pelas duas cópias, o que é chamado de haploinsuficiência. As deleções podem ser de três tipos: *terminais*, quando acontecem por causa de uma quebra simples na extremidade de um cromossomo; *intersticiais*, quando resultam de uma quebra dupla ao longo do braço de um cromossomo, causando ligação dos segmentos adjacentes e a perda do segmento interno às rupturas; e *dupla terminal*, que consiste em rupturas terminais nas duas extremidades de um cromossomo, com perda dos segmentos finais acêntricos. Essa alteração leva à formação de extremidades reativas capazes de gerar um cromossomo em formato circular, denominado *cromossomo em anel*.

Deleções cromossômicas
Perdas de um segmento cromossômico em decorrência de quebras transversais

Haploinsuficiência
Quando a presença de apenas uma cópia normal de um alelo não é suficiente para prevenir a manifestação de um fenótipo anormal em virtude da perda de função do outro alelo

Um exemplo clássico de deleção cromossômica levando a uma cromossomopatia é a *síndrome de cri-du-chat* (46,XX,5p- ou 46,XY,5p-), uma disfunção congênita rara (1 a cada 50.000 nascimentos), causada pela *deleção parcial do braço curto do cromossomo 5*. A sín-

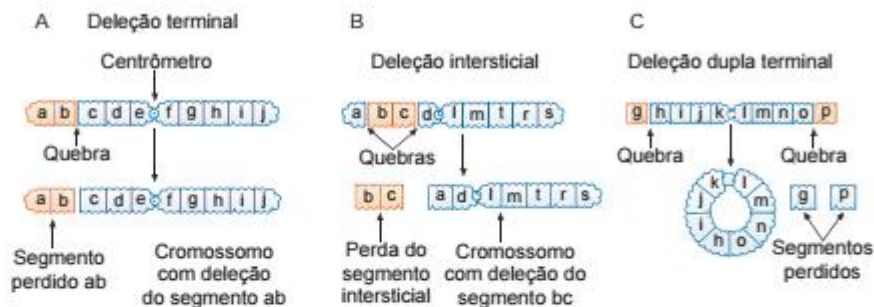



Figura 4.16 Principais tipos de deleção: (A) deleção terminal, (B) deleção intersticial, (C) deleção dupla terminal.

drome recebe esse nome pelo fato de os bebês apresentarem um choro característico, semelhante ao miado de um gato. As principais manifestações clínicas da síndrome incluem ainda microcefalia, pregas epicânticas, orelhas de baixa implantação, hipotonia, hipertelorismo e deficiência intelectual de moderada a grave. A gravidade do fenótipo depende da extensão da deleção, mas a região crítica, em haploinsuficiência em todos os portadores da síndrome, parece estar na região 5p15.

Além das quebras transversais, as deleções podem surgir também de uma *recombinação cromossômica desigual* durante a meiose, em virtude de um alinhamento imperfeito entre os cromossomos homólogos ou por meio da formação de estruturas em forma de grampo, durante a replicação do DNA. Esses eventos ocorrem principalmente quando no ponto de recombinação e/ou replicação estão envolvidas sequências consecutivas de DNA repetitivo, podendo levar também à formação de *duplicações* do material cromossômico, como podemos ver na Figura 4.17.

Aneussomia segmentar
Perda de um segmento cromossômico em um dos membros de um par de homólogos, levando à haploinsuficiência de genes contidos no segmento cromossômico perdido

Recombinação homóloga não alélica
Recombinação entre sequências homólogas desalinhadas que gera duplicações e deleções

As deleções podem provocar efeitos fenotípicos diversos com a perda de genes contíguos múltiplos na região da deleção, as denominadas *síndromes de genes contíguos*. A haploinsuficiência dos genes presentes nesse segmento cromossômico que foi perdido em um dos cromossomos homólogos é denominada aneussomia segmentar. A análise molecular mais refinada em indivíduos que apresentam a mesma síndrome de genes contíguos tem revelado a presença de sequências repetitivas com pequeno número de cópias próximas à região deletada. Essas repetições podem favorecer o aparecimento desses rearranjos por recombinação homóloga não alélica, assim como podem desencadear a presença de duplicações. Na Tabela 4.2 estão representadas as principais síndromes de genes contíguos.

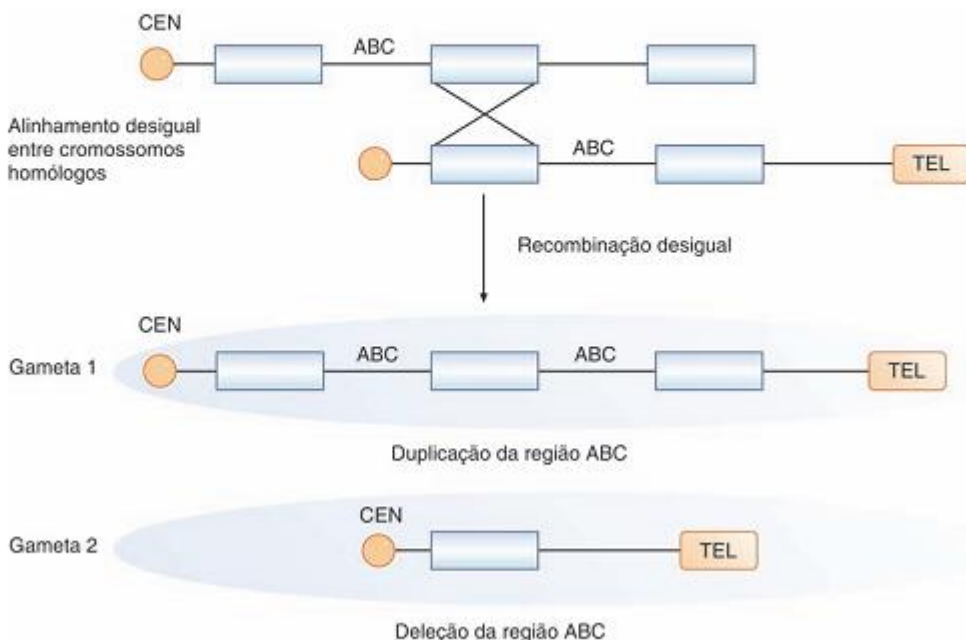



Figura 4.17 Recombinação desigual durante a formação dos gametas pode causar deleções e duplicações. Centrômero (Cen); telômero (Tel).

■ **Tabela 4.2** Algumas das síndromes de genes contíguos, sua localização cromossômica e principais características fenotípicas associadas.

Síndrome	Localização citogenética	Principais características fenotípicas
Deleção 1p36	1p36	Atraso do desenvolvimento neuropsicomotor e/ou deficiência intelectual, hipotonia, convulsões, déficit auditivo, retardo de crescimento e/ou obesidade; microcefalia, braquicefalia, dismorfias faciais
Sotos	5q35	Crescimento excessivo nos primeiros anos de vida, deficiência intelectual, hipotonia, macrocefalia, dismorfias faciais típicas
Williams	7q11.23	Face característica, hipercalcemia, problemas cardiovasculares, personalidade sociável, deficiência intelectual
Langer-Giedion	8q24.1	Atraso no desenvolvimento, cabelos finos, nariz bulboso, exostoses múltiplas, epífises em forma de cone
WAGR	11p13	Tumor de Wilms, aniridia, displasia geniturinária, retardo do crescimento, deficiência intelectual
Angelman	15q11-13	Deficiência intelectual, risos sem motivo, microcefalia, convulsão, ausência de fala, problemas de movimento e equilíbrio
Prader Willi	15q11-13	Obesidade, hipotonia, deficiência intelectual, hiperfagia, mãos e pés pequenos, hipogonadismo, baixa estatura
Charcot-Marie Tooth	17p12	Perda de tecido muscular e sentido de toque, predominantemente nos pés e pernas
Smith-Magenis	17p11.2	Deficiência intelectual, alterações congênitas múltiplas, distúrbios do sono e problemas comportamentais
Miller-Diecker	17p13.3	Lisencefalia, deficiência intelectual, face característica
Velo-cardiofacial	22q11	Deficiência intelectual, fissura palatina, face típica, anomalias cardíacas

Duplicação cromossômica
Alteração estrutural que envolve a adição de uma cópia de um determinado segmento cromossômico

As **duplicações cromossômicas** representam repetições de um segmento cromossômico, tendo como consequência o aumento do número de genes, originalmente presentes na célula. Essas alterações têm grande *importância evolutiva*, tendo em vista que ao longo do tempo uma sequência duplicada pode adquirir novas funções gênicas, diferentes daquelas atribuídas às suas sequências originais, de modo a aumentar a variabilidade da espécie. Como ilustrado na Figura 4.18, as duplicações podem ser classificadas em: *duplicações em tandem*, quando o segmento duplicado foi inserido na mesma posição e orientação em que se encontrava no segmento original; *em tandem invertida*, quando o segmento duplicado foi reintroduzido na mesma posição de origem de forma invertida; e *duplicação por segmento deslocado*, quando o segmento duplicado é inserido em um cromossomo diferente daquele em que se encontrava originalmente.

Em geral, os efeitos fenotípicos das duplicações não são tão graves como aqueles provocados pelas deleções, mas a alteração da expressão de genes no ponto da duplicação pode resultar em anomalias importantes. Um exemplo de duplicação levando a um fenótipo anormal é a raríssima síndrome do olho de gato, na qual ocorre uma duplicação intersticial da região

Figura 4.18 Diferentes tipos de duplicações cromossômicas.

22q11, causando coloboma inferior da íris, atresia anal, hipertelorismo, deficiência intelectual de graus variáveis e defeitos cardíacos em mais de 1/3 dos casos.

As *inversões* constituem outra categoria de alterações cromossômicas estruturais intracromossômicas representadas por duas quebras transversais ao longo de um cromossomo, seguidas da reintrodução do segmento clivado na mesma posição, mas em orientação invertida. Como podemos observar na Figura 4.19, as inversões podem ser *pericêntricas*, quando as rupturas acontecem em cada braço cromossômico e o *segmento invertido inclui o centrômero* ou *paracêntricas*, quando ambas as rupturas ocorrem em um mesmo braço do cromossomo, *não envolvendo o centrômero*. Em geral, por se tratar de um rearranjo balanceado, sem adição ou perda de material genético, o portador de uma inversão não exibe um fenótipo anormal, a menos que genes importantes nos pontos de quebras tenham sido afetados. No entanto, a importância clínica da inversão está no fato do portador de uma *inversão balanceada* poder produzir uma *progênie com fenótipo anormal*. Isso ocorre em virtude do pareamento incorreto do cromossomo portador da inversão e seu homólogo normal durante a meiose I, pela formação de uma alça na região da inversão e uma consequente propensão à formação de gametas desbalanceados. Durante o *crossing over*, uma inversão paracêntrica pode levar tanto à formação de cromossomos *dicêntricos* (com dois centrômeros) como de cromossomos acêntricos, enquanto a inversão pericêntrica pode gerar cromossomos com deleções e duplicações.

Inversão
Alteração que envolve a excisão de um segmento cromossômico, seguida de um movimento de 180° e a reintrodução do segmento em ordem invertida. Pode ser paracêntrica ou pericêntrica

A principal inversão conhecida em humanos é a inversão pericêntrica do cromossomo 9 [inv(9)(p11q12)], presente em 1% dos indivíduos analisados citogeneticamente. Essa inversão, em geral, não causa danos a seus portadores, mas algumas evidências apontam para um risco aumentado de infertilidade em homens com esta inversão em virtude da formação das alças ou de fragmentos acêntricos durante a espermatogênese.

Os *isocromossomos* constituem outro subgrupo de alterações estruturais, formado quando a divisão do centrômero e a separação das cromátides-irmãs, durante a divisão celular, ocorre de modo *transversal* em vez de longitudinal. Com isso, são originados dois cromossomos, cujos braços são iguais. Um exemplo de isocromossomo associado a uma entidade clínica é a síndrome de Turner por formação de isocromossomo de braço longo do cromossomo X [46,X,i(Xq)] e haploinsuficiência de genes importantes no cromossomo X.

Alterações estruturais mais raras incluem ainda a formação de cromossomos que apresentam dois centrômeros (*dicêntricos*) e de *cromossomos marcadores*. Os cromossomos *dicêntricos* resultam da *fusão de dois segmentos cromossômicos*, cada qual com seu centrômero, provenientes de cromossomos diferentes ou de duas cromátides de um único cromossomo, com perda subsequente dos fragmentos acêntricos. Os cromossomos dicêntricos poderão ser estáveis durante a divisão celular se um dos centrômeros estiver inativado. Já os *cromossomos marcadores* ou supranumerários são cromossomos muito pequenos, estruturalmente anormais, que podem ser observados em preparações citogenéticas, como um complemento ao conteúdo cromossômico básico, geralmente observado em estado de mosaicismo. O material extra presente nesses

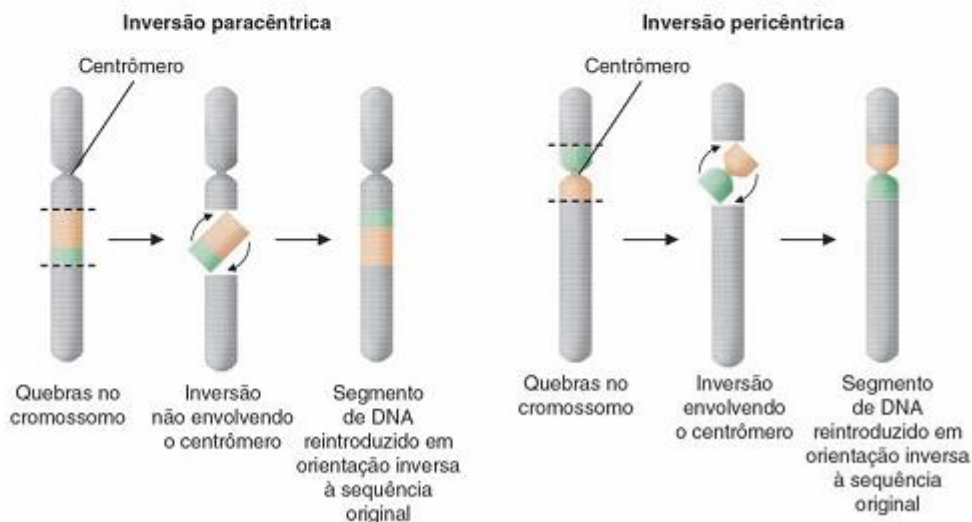



Figura 4.19 Os dois tipos de inversão cromossômica (paracêntrica e pericêntrica).

cromossomos tem heterocromatina centromérica, que garante certa estabilidade ao longo das divisões celulares. Todavia, o tamanho diminuto traz grandes dificuldades para a caracterização desses cromossomos, por meio de técnicas de citogenética convencionais. Dependendo da porcentagem de células contendo o cromossomo marcador e da constituição gênica deste, o portador desse tipo de alteração pode ter ou não alguma consequência fenotípica.

É válido ressaltar que alguns cromossomos em particular têm pequenas variações intracromossômicas em sua estrutura ou tamanho sem que estas representem consequências fenotípicas para os indivíduos que as têm. Exemplos de variações desse tipo incluem as regiões de heterocromatina presentes nos cromossomos 1, 9, 16 e Y, como também o tamanho dos satélites e hastes presentes nos braços curtos dos cromossomos acrocêntricos (13, 14, 15, 21 e 22).

Alterações estruturais intercromossômicas

Até o momento, descrevemos as principais alterações cromossômicas que podem ocorrer dentro de um mesmo cromossomo e suas consequências fenotípicas. Agora, detalharemos as alterações cromossômicas estruturais que envolvem mais de um cromossomo. As *alterações intercromossômicas* são principalmente representadas pelas *translocações*, que envolvem quebras em *cromossomos diferentes* (não homólogos), seguidas na maioria das vezes da *permuta recíproca* dos segmentos quebrados entre eles. Em geral, somente dois cromossomos estão envolvidos em uma translocação e, como a troca é bidirecional, e o número de cromossomos na célula permanece idêntico ao original (46 cromossomos). Quando não há perda de segmentos acêntricos em uma translocação, a alteração será balanceada e muitas vezes não causa danos fenotípicos ao seu portador. Contudo, durante a gametogênese, o pareamento dos cromossomos translocados envolve a formação de uma *estrutura quadrivalente* (em formato de cruz), possibilitando que durante a anáfase os cromossomos segreguem formando seis tipos diferentes de gametas. Destes, um gameta será normal, um apresentará a mesma translocação recíproca balanceada e os quatro restantes serão desbalanceados. Observe a Figura 4.20.

Um tipo peculiar de *translocação* é restrito aos *cromossomos acrocêntricos* (13, 14, 15, 21 e 22) e, por ter sido descrita por Robertson em 1919, recebe a denominação translocação robertsoniana (rob). Nesse grupo especial de translocações, dois cromossomos acrocêntricos se fundem unindo os dois braços longos dos cromossomos envolvidos com perda dos braços curtos. O indivíduo portador desse tipo de translocação terá somente 45 cromossomos, mas seu fenótipo será balanceado. Isso acontece porque o material genético que estava presente nos braços curtos e é perdido contém somente genes para a formação de RNA ribossômico, que estão redundantemente presentes nos demais cromossomos acrocênticos, o que torna possível a neutralização dessa carência.

Translocação robertsoniana
Tipo especial de translocação dos cromossomos acrocêntricos, na qual ocorre fusão no centrômero ou próximo a ele, com perda subsequente dos braços curtos

Um exemplo de translocação robertsoniana é a que envolve um cromossomo 21 e um 14 em um indivíduo que terá, portanto, 45 cromossomos em suas células. Embora seja fenotipicamente normal, no momento de formar seus gametas, assim como acontece na translocação recíproca, serão formados seis tipos diferentes de gametas, conforme ilustrado na Figura 4.21. Caso haja fecundação por um gameta normal, teremos:

- 1/6 de indivíduos normais com 46 cromossomos
- 1/6 de portadores da mesma translocação balanceada presente no genitor [45,XX ou 45,XY,der(14;21)(q10;q10)]
- 1/6 de indivíduos com *síndrome de Down* por *translocação robertsoniana*. Nesse caso, embora os indivíduos apresentem 46 cromossomos, existem três cópias do cromossomo 21 (duas cópias livres do cromossomo 21 somadas a uma terceira cópia translocada a um dos cromossomos 14). O cariótipo, nesse caso será representado como 46,XX ou 46,XY,der(14;21)(q10;q10),+21
- 1/6 de zigotos com monossomia do cromossomo 21 (45 cromossomos no total)
- 1/6 dos indivíduos com monossomia do cromossomo 14 (45 cromossomos no total)
- 1/6 dos zigotos com trissomia do cromossomo 14 por translocação robertsoniana. Nesse caso, embora os indivíduos apresentem 46 cromossomos, existem três cópias do cromossomo 14 (duas cópias livres do cromossomo 14 somadas a uma terceira cópia translocada a um dos cromossomos 21).

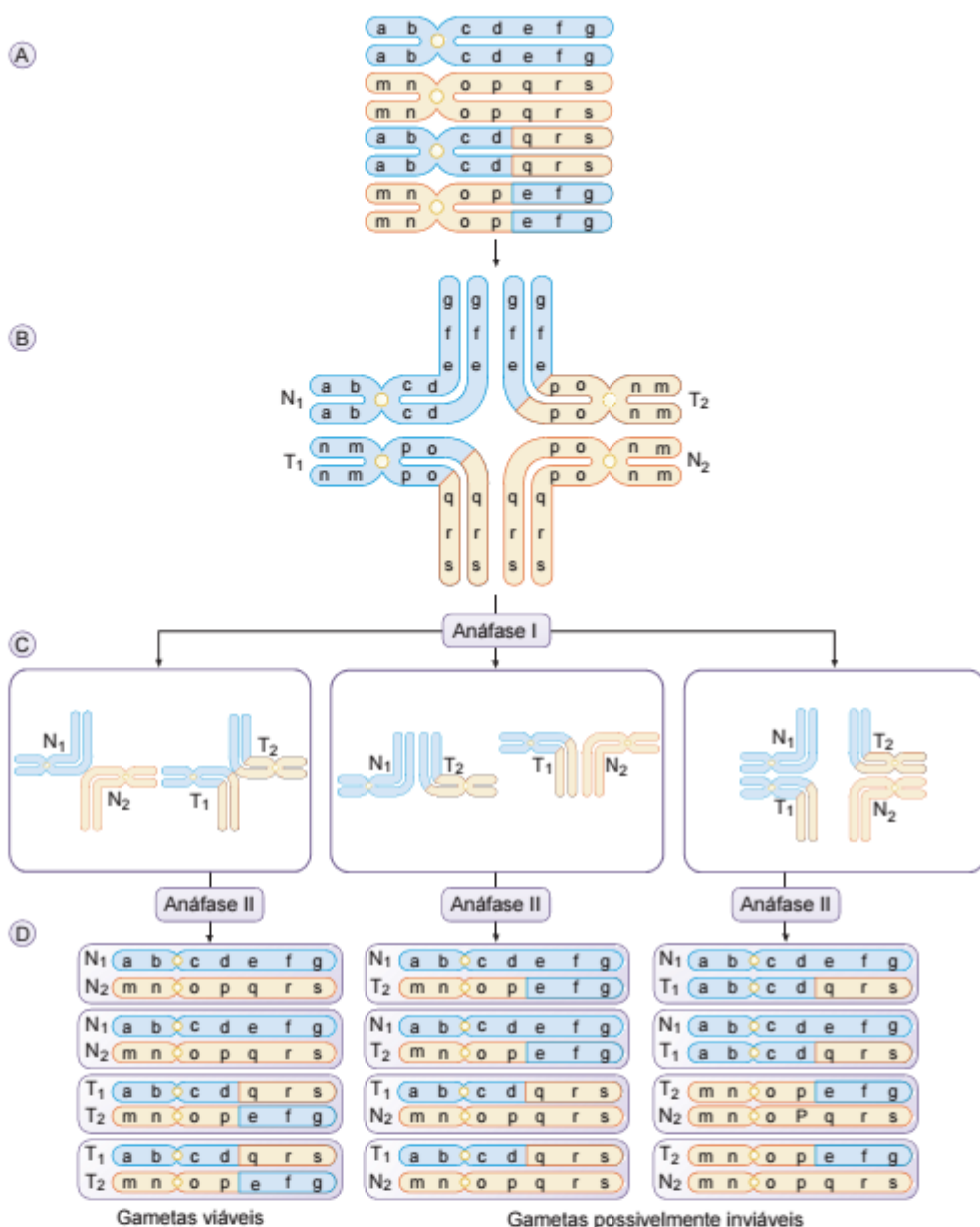



Figura 4.20 Diagrama de uma translocação balanceada recíproca levando à formação de gametas viáveis e não viáveis. (A) Indivíduo com uma translocação recíproca balanceada; (B) pareamento dos cromossomos homólogos envolvidos em uma translocação recíproca por meio de uma estrutura quadrirradial formada durante a prófase I da meiose para a formação de gametas; (C) possibilidades durante a segregação dos cromossomos na anáfase I; (D) constituição final dos gametas (viáveis e possivelmente inviáveis) após a anáfase II.

Considerando a *viabilidade* dos zigotos, visto que as possibilidades (d), (e) e (f) são incompatíveis com a vida, um portador de uma translocação robertsoniana balanceada 14q21q que case com um parceiro normal terá 1/3 de chance de ter um filho com síndrome de Down, 1/3 de chance de ter filhos normais e 1/3 de chance de ter filhos portadores da translocação balanceada.

Risco de recorrência
Risco de uma condição aparecer novamente em uma mesma família

Com isso, o risco de um casal que já teve uma criança com síndrome de Down ter outra criança com a mesma síndrome, denominado de risco de recorrência, pode variar entre as formas causadas por não disjunção dos casos de translocação robertsoniana. Nos casos de não disjunção cromossômica, o risco de se ter outra criança com trissomia do 21 é o mesmo para a população em geral (1%), mas caso uma translocação robertsoniana entre os cromossomos 14 e 21 tenha sido identificada de maneira balanceada em um dos genitores, o risco sobe para 1/3. Caso os cromossomos envolvidos na translocação robertsoniana balanceada sejam os dois cromossomos 21, esse risco aumentará para 100% dentre os indivíduos viáveis, tendo em vista que a monossomia do 21 é letal. Sendo assim, embora a síndrome de Down gere uma condição fenotípica facilmente identificável (até mesmo por leigos), a importância da identificação do cariótipo relacionado com sua manifestação está no aconselhamento genético dos casais que pretendem ter mais filhos. Além disso, a caracterização do

Figura 4.21 Representação esquemática de como uma translocação robertsoniana balanceada entre os cromossomos 14 e 21 pode levar à formação de gametas anormais. (A) quebra em braços opostos dos cromossomos 14 e 21 com fusão dos braços longos e perdas dos braços curtos; (B) tipos de gametas produzidos por um portador da translocação balanceada entre os cromossomos 14 e 21; (C) tipos de zigotos formados a partir da fecundação dos gametas demonstrados em (B) por um gameta cromossomicamente normal.

tipo de alteração cromossômica relacionada com o cromossomo 21 extra é importante para o prognóstico dos indivíduos com síndrome de Down, tendo em vista que portadores de cariótipos atípicos parecem estar mais propensos ao desenvolvimento de doenças como a leucemia, por exemplo.

Além das translocações recíprocas, outro tipo de alteração intercromossômica que pode ser citada é a *inserção* cromossômica (ins). As inserções cromossômicas representam um *evento unidirecional*, no qual ocorre incorporação de um grande segmento de um cromossomo em um segundo cromossomo, sem que haja troca entre eles. Para que esse evento aconteça, é necessário que haja *três quebras cromossômicas* (duas rupturas para liberar o fragmento que será inserido e uma ruptura no segundo cromossomo para que o fragmento possa se inserir). Inserções também podem propiciar o desbalanceamento do rearranjo durante a gametogênese.

■ Cromossomos X e Y

Em nível citogenético, o cromossomo X é consideravelmente maior que o cromossomo Y, o que é refletido também em nível molecular, visto que o cromossomo X apresenta

mais de 1180 genes funcionais descritos no *Ensembl*, contra 97 genes que se expressam no cromossomo Y. No entanto, os cromossomos sexuais X e Y foram originados de *um par de autossomos ancestrais* há cerca de 300 milhões de anos. Ao longo do processo evolutivo, uma barreira parcial para a recombinação foi estabelecida entre esses dois cromossomos em virtude do surgimento de inversões e, embora muitos elementos funcionais originais tenham sido conservados no cromossomo X, o cromossomo Y degenerou, perdendo grande parte dos traços do autossomo ancestral. O par de cromossomos sexuais, diferentemente dos demais cromossomos, passou então a ter características e mecanismos de regulação genética próprios e bem distintos entre si. Com isso, foram criados dois sexos diferentes, o sexo homogamético, no qual os dois cromossomos sexuais são iguais (46,XX) e podem recombinar como qualquer par de autossomos durante a gametogênese e o sexo heterogamético, representado pelos homens, que apresentam um cromossomo X e um Y. Nos homens, portanto, o emparelhamento e a recombinação meiótica envolverão apenas as extremidades desses cromossomos, que mantiveram um alto grau de similaridade entre os cromossomos X e Y ao longo do processo evolutivo. Essas regiões homólogas são chamadas de regiões pseudoautossômicas e são divididas em PAR1 e PAR2. Os genes presentes nas regiões pseudoautossômicas codificam proteínas necessárias a ambos os sexos, participando da transdução de sinal, crescimento ósseo, metabolismo energético e síntese de hormônios e receptores.

Sexo homogamético
Sexo representado pela presença de dois cromossomos sexuais iguais, ou seja, uma mulher (46,XX)

Sexo heterogamético
Sexo representado pela presença de dois cromossomos sexuais distintos, ou seja, um homem (46,XY)

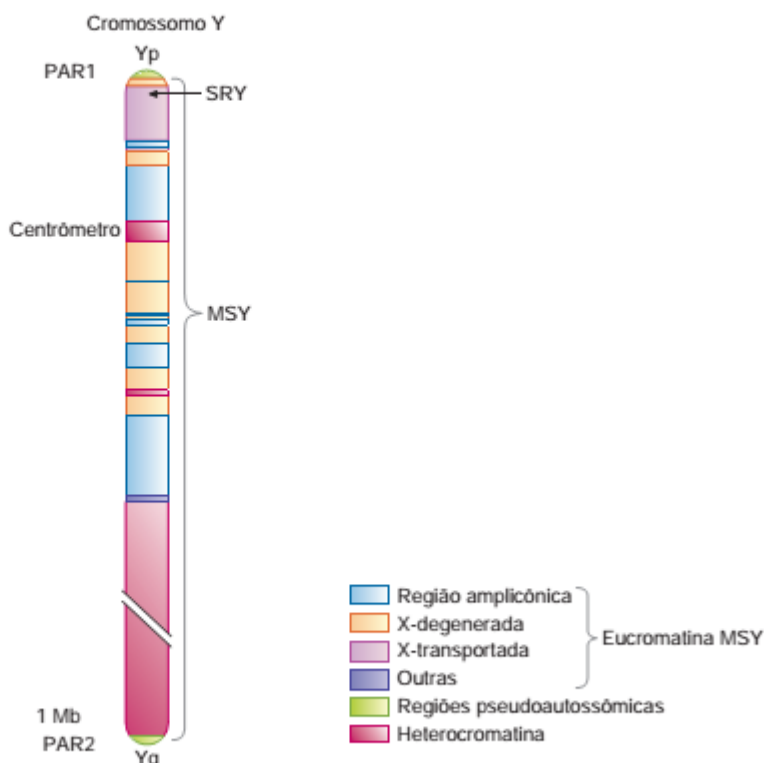
Região pseudoautossômica
Porção final dos braços curtos e longos dos cromossomos X e Y que têm homologia e que trocam material genético entre si durante o *crossing over* que ocorre na formação dos gametas masculinos

A região pseudoautossômica de maior extensão (PAR1) tem 2,6 Mb e localiza-se na área terminal dos braços curtos do X e do Y, enquanto a PAR2 apenas 320 kb, e encontra-se na extremidade oposta dos braços longos. Os cerca de 24 genes contidos na região pseudoautossômica PAR1 têm alta taxa de recombinação e não estão sujeitos à compensação de dose regida pela inativação de um dos cromossomos X na mulher (Capítulo 2). Já os 5 genes presentes em PAR2 foram criados em virtude da duplicação de material genético do cromossomo X para o Y ao longo da evolução. Parte desses genes pode sofrer inativação, provavelmente como reflexo do *status* do cromossomo X, antes do evento de duplicação.

Como demonstrado na Figura 4.22, a área central do cromossomo Y entre as regiões PAR1 e PAR2 não se recombina com o cromossomo X e é denominada *male-specific region* (MSY), compreendendo 95% de todo o cromossomo. Essa região específica do sexo masculino representa uma mistura de *regiões de heterocromatina* e 3 classes diferentes de *sequências de eucromatina*: a *X-transposta*, a *X-degenerada* e a *amplicônica*. Com relação à heterocromatina, o cromossomo Y apresenta duas principais regiões desse tipo: uma próxima ao centrômero, como é característico para um cromossomo nuclear, e um grande bloco de *heterocromatina*, contendo DNA altamente repetitivo que se estende por uma vasta área de seu braço longo. Já a região *X-transposta* é formada por sequências que exibem similaridade com o cromossomo X e vieram transpostas da região Xq21 para o cromossomo Y há cerca de 3 a 4 milhões de anos. A região *X-degenerada*, por sua vez, corresponde a sequências de DNA remanescentes do cromossomo autossômico ancestral. Por último, a *região amplicônica* representa cópias internas de sequências no cromossomo Y que têm alta similaridade entre si. A maioria dos genes codificadores de proteínas nos segmentos amplicônicos se expressa especificamente nos testículos e está provavelmente envolvida no desenvolvimento masculino. Além da recombinação homóloga com o cromossomo X que ocorre nas regiões pseudoautossômicas, a região amplicônica do cromossomo Y apresenta um mecanismo próprio de recombinação intracromossômica, chamado de conversão gênica. Este representa a transferência não recíproca da informação da sequência de um duplex de DNA para outro, ocorrendo entre cópias de sequências internas específicas do cromossomo Y. Nesse processo natural, uma sequência é alterada para se tornar idêntica à sequência que foi utilizada como molde. Acredita-se que a conversão gênica tenha sido a maneira encontrada pelo cromossomo Y para preservar a estrutura e funções de genes ativos, principalmente as testiculares, da atuação de forças evolutivas na ausência de *crossing over*.

Conversão gênica
Permuta genética não recíproca natural, na qual uma sequência de DNA é modificada para se tornar idêntica à sequência que foi utilizada como molde (doadora). É um processo observado no cromossomo Y para preservar a função de seus genes ao longo da evolução

Dentre os genes funcionais localizados no cromossomo Y, um merece atenção especial: o *gene determinante sexual do cromossomo Y* (*SRY*), o qual está localizado no braço curto do

Figura 4.22 Ideograma representativo do cromossomo Y. Note as regiões pseudoautossômicas (PAR1 E PAR2) nas extremidades do cromossomo e o detalhe do gene SRY localizado logo abaixo da região pseudoautossômica PAR1.

cromossomo Y (Yp11.31), logo abaixo da região pseudoautossômica. Esse gene codifica um *fator de transcrição* que regula outros genes e é o principal componente envolvido na *determinação sexual* e na *formação dos testículos*. Além do SRY, vários genes relacionados com a espermatogênese também têm sido descritos no cromossomo Y. Na região de eucromatina Xq11, foram localizados genes para *fatores de azoospermia* (AZF) em três regiões distintas (AZFa, AZFb e AZFc). Deleções e mutações pontuais nestas regiões podem afetar um ou mais genes nelas contidas, levando à azoospermia.

Ao mesmo tempo em que a evolução dos cromossomos sexuais preservou genes relacionados com a masculinidade no cromossomo Y, ela foi acompanhada de um enriquecimento desproporcional de genes envolvidos no desenvolvimento e na função cerebral no cromossomo X. Isso ocorre seja pelo recrutamento de novos genes, seja pela aquisição de uma nova função para aqueles genes anciães presentes no arcabouço do cromossomo X. De todos os genes codificadores de proteínas presentes no cromossomo X listados no *Ensembl*, cerca de 40% são expressos no cérebro. O acúmulo de genes relacionados com a função cognitiva no cromossomo X em relação aos autossomos pode ter sido resultado da rápida fixação de alelos recessivos que conferiam vantagens aos machos hemizigotos. Essa evidência tem levado alguns autores a acreditarem que a seleção natural direcionou o acúmulo de genes moduladores da inteligência no cromossomo X, sugerindo que tanto a deficiência intelectual quanto a inteligência excepcional poderiam ser mais frequentes em homens que só tem um cromossomo X. Curiosamente, mulheres portadoras da síndrome de Turner (45,X), cujo cromossomo X único tem herança paterna apresentam uma melhor habilidade social-cognitiva. Além disso, testes de imagem em encéfalos de portadores das síndromes de Turner (45,X) e Klinefelter (47,XXY) mostram diferenças significativas na integridade neuronal e no volume de massa cinzenta, de acordo com a origem parental dos cromossomos X, com maior habilidade verbal e comportamental naqueles indivíduos com cromossomos X de origem paterna. Outro indício de que o cromossomo X reúne genes associados à função intelectual vem do estudo de gêmeos. Gêmeas idênticas apresentam a mesma configuração genética, com dois cromossomos X, dos quais um será inativado pelo *mecanismo de compensação de dose*. No entanto, como a *inativação* é

Hemizigoto
Termo utilizado principalmente para os genes no cromossomo X que estão apenas em uma cópia nos homens em relação às mulheres

aleatória (tanto o cromossomo X materno quanto o paterno poderão estar inativos), o padrão de inativação do cromossomo X pode não ser exatamente o mesmo para todos os diversos tecidos. Sendo assim, as irmãs podem expressar genes localizados no cromossomo X em proporções distintas em um dado tecido. Em contraste, gêmeos masculinos idênticos herdaram somente um cromossomo X de suas mães, o qual sempre estará sendo expresso da mesma maneira nos diferentes tecidos. Com isso, em alguns tratos relacionados com a inteligência, como as habilidades verbais e um bom comportamento social, os gêmeos idênticos homens são mais parecidos entre si do que as gêmeas mulheres.

Alterações cromossômicas relacionadas com os cromossomos sexuais

As alterações envolvendo os cromossomos sexuais também podem ser numéricas ou estruturais, alcançando todas as células de um indivíduo ou parte delas (mosaicismo), como visto para os autossomos. No entanto, com exceção à monossomia do cromossomo Y, todas as aneuploidias envolvendo cromossomos sexuais (inclusive tetra e pentassomias do cromossomo X) são compatíveis com a vida e podem estar presentes em indivíduos com desenvolvimento normal ou próximo do normal. Uma justificativa para esta observação é a inativação *epigenética* do cromossomo X, que acontece em todos os cromossomos X adicionais.

Para não haver desequilíbrio na expressão de genes entre os sexos feminino e masculino, em decorrência do tamanho discrepante entre os dois cromossomos, na fase inicial do desenvolvimento embrionário, um dos cromossomos X da mulher é inativado (silenciado). A esse mecanismo dá-se o nome de *compensação de dose* e, com isso, a mulher 46,XX terá em suas células somáticas sempre um cromossomo X ativo e outro inativo, enquanto o único cromossomo X presente no homem 46,XY estará sempre ativo. A inativação ocorrerá como consequência de um complexo mecanismo que envolve diferentes fatores epigenéticos (detalhadamente descritos no Capítulo 2) responsáveis por desligar os genes presentes em um dos cromossomos X.

A escolha por qual cromossomo X será inativado na mulher (paterno ou materno) é aleatória, mas o mesmo padrão é mantido nas células-filhas descendentes. Com isso, as mulheres representam mosaicos para a expressão de genes do cromossomo X, visto que algumas células do organismo expressarão os genes oriundos do cromossomo X paterno, enquanto outras expressarão o X materno. O cromossomo X inativo tem a característica de apresentar-se como uma massa heterocromática de fácil visualização durante a intérfase, recebendo o nome de corpúsculo de Baar nos preparos tradicionais de lâminas para visualização da cromatina sexual. Essa metodologia utiliza como princípio a característica do cromossomo X inativo de se replicar tardiamente durante a interfase em comparação aos demais cromossomos. Por isso, nessa fase, o cromossomo X inativo apresenta-se mais compactado e é, portanto, mais facilmente corado. A heterocromatina do cromossomo X inativo, diferentemente da heterocromatina constitutiva representada principalmente pelos centrômeros e pelas constrições secundárias, é de natureza facultativa, pois o cromossomo X que está inativo em algumas células aparecerá como eucromatina em outras. Como pode ser visto na Tabela 4.3, homens 46,XY não terão

Corpúsculo de Baar
Estrutura condensada, altamente corável, que é encontrada na maioria das células de mamíferos e representa o cromossomo X inativo

Heterocromatina constitutiva
Cromatina que aparece sempre de forma heterocromática, presente nos centrômeros

Heterocromatina facultativa
Cromatina que pode existir como heterocromatina ou eucromatina, dependendo do estado da célula

Tabela 4.3 Número de corpúsculos de Baar em diferentes complementos de cromossomos sexuais.

Constituição cromossômica sexual	Síndrome	Número de corpúsculos de Baar
XX	Mulher normal	1
XY	Homem normal	0
X	Turner	0
XXX	Triplo X	2
XXXX	Tetrassomia do X	3
XXY	Klinefelter	1
XXXY	Klinefelter variante	2
XYY	Duplo Y	0

corpúsculo de Baar, enquanto a mulher 46,XX exibirá apenas um corpúsculo. Em alterações numéricas dos cromossomos sexuais, qualquer cromossomo X extra que surja será inativo, ao passo que na monossomia do X (45,X), o único cromossomo X presente estará sempre ativo. Além disso, em alterações que envolvem o cromossomo X, há em geral inativação preferencial do cromossomo X anormal.

Atualmente, com os avanços no conhecimento de fatores epigenéticos que diferenciam o cromossomo X inativo do ativo, alguns marcadores passaram a ser utilizados na detecção de alterações numéricas associadas ao cromossomo X no núcleo interfásico. Um desses marcadores é representado pela variante de histona macroH2A, que é preferencialmente incorporada na cromatina do cromossomo X inativo.

Características fenotípicas ligadas a alterações em cromossomos sexuais não são tão graves quanto em autossomos, mas podem afetar a determinação e a diferenciação sexuais e a infertilidade ou gerar características fenotípicas mais generalizadas.

Além disso, as aneuploidias dos cromossomos sexuais são mais frequentes que as dos autossomos, com uma frequência geral de 1 a cada 500 nascimentos. Isso pode ser explicado pelo fato de as aneuploidias autossômicas estarem mais suscetíveis à seleção natural ainda no desenvolvimento fetal ou por maior propensão dos cromossomos sexuais à não disjunção meiótica em relação aos autossomos. Nesse último aspecto, devemos considerar que, particularmente na meiose masculina, os cromossomos X e Y são pareados apenas por meio das regiões pseudoautossômicas localizadas em suas extremidades, e a sinapse é, portanto, *menos estável*, o que pode favorecer erros na segregação cromossômica. Com relação às divisões mitóticas, a incidência de mosaicismo envolvendo os cromossomos sexuais também é maior em relação aos autossomos, além de representar um percentual significativo (10% a 20%) dentre todas as aneuploidias de cromossomos sexuais. Devemos ter em mente, no entanto, que a presença de anomalias no desenvolvimento sexual não necessariamente é causada por alterações dos cromossomos sexuais perceptíveis ao cariótipo. Essas condições podem resultar de alterações em um ou poucos genes ou de causas não genéticas. A identificação dessas alterações é o primeiro passo para o reconhecimento dos genes implicados na determinação e na diferenciação sexual.

Quatro principais síndromes relacionadas com os cromossomos sexuais são reconhecidas citogeneticamente, conforme a Tabela 4.4 a seguir.

■ Tabela 4.4 Características fenotípicas, cariótipos associados e incidência aproximada das principais alterações cromossômicas sexuais.

Síndrome	Características fenotípicas	Cariótipos associados	Incidência aproximada
Síndrome de Turner	Baixa estatura, pescoço alado, atraso na maturação sexual, gônadas em fita, dentre outras	45,X 46,X,i(Xq) Outros (deleções, mosaicismo, cromossomos em anel)	1/5.000 mulheres 1/50.000 mulheres 1/15.000 mulheres
Síndrome de Klinefelter	Estatura alta; hipogonadismo; infertilidade	47,XXY 48,XXXY Outros (48,XXYY; 49,XXXY; mosaicos; 46,XX com translocação do gene SRY para o cromossomo X)	1/1.000 homens 1/25.000 homens 1/15.000 homens
Síndrome do duplo Y	Estatura alta	47,XYY	1/1.000 homens
Síndrome do triplo X	Estatura alta	47,XXX	1/1.000 mulheres

Exemplos de alterações cromossômicas sexuais

Síndrome de Turner (45,X e demais cariótipos atípicos)

A *síndrome de Turner* é uma condição que acomete 1 a cada 5.000 recém-nascidos do sexo feminino. Ainda que tradicionalmente a síndrome seja atribuída ao cariótipo 45,X, sabe-se que em 50% dos casos, a condição surge em decorrência de outras alterações numéricas ou estruturais mais atípicas que levam à perda parcial ou total do segundo cromossomo X na mulher,

como podemos observar na Tabela 4.5. Essas alterações incluem a formação de isocromossomo do braço longo do cromossomo X, cromossomos em anel ou demais deleções que provocam a haploinsuficiência de genes importantes no braço curto do cromossomo X.

■ **Tabela 4.5** Constituições cromossômicas relacionadas com a síndrome de Turner.

Constituição cromossômica	Proporção entre os casos de síndrome de Turner
45,X	50%
46,X,i(Xq)	15%
45,X/46,XX	15%
45,X/46,X,i(Xq)	aproximadamente 5%
45,X, outra anormalidade do cromossomo X	aproximadamente 5%
Outra 45,X/?	aproximadamente 5%

O cariótipo 45,X é considerado a alteração mais comum em abortos espontâneos. Com isto, acredita-se que o cariótipo 45,X gere uma sobrevida intrauterina diminuída, como as demais monossomias e que os indivíduos que nascem com a síndrome representariam casos de mosaicismo somático não diagnosticados envolvendo uma linhagem normal. Tanto em abortos como em nativivos, em aproximadamente 80% dos casos, o cromossomo X único é de origem materna, de modo que a não disjunção que origina a alteração aconteceu na meiose paterna. Dentre os casos de mosaicismo somático identificados, a linhagem 45,X pode estar acompanhada de uma ou mais linhagens celulares com um cromossomo X completo (p. ex., 45,X/46,XX/47,XXX) ou estruturalmente anormal (p. ex., 45,X/46,X,i(Xq)). No primeiro caso, a presença de duas ou mais linhagens diferentes que envolvem o cromossomo X é proveniente de não disjunção após a formação do zigoto, enquanto no segundo, os erros mitóticos foram precedidos de erros meióticos.

Entre as aneuploidias relacionadas com os cromossomos sexuais, as mulheres com síndrome de Turner são as que apresentam um quadro fenotípico mais perceptível ao nascimento ou antes da puberdade. No entanto, achados sugestivos da síndrome podem ser observados ainda no desenvolvimento pré-natal por meio do exame de ultrassonografia e incluem higroma cístico, hidropsia fetal, edema subcutâneo, fêmur curto, aumento da translucência nucal, além de anormalidades cardíacas e renais. Ao nascimento, as características mais comuns são o pescoço alado, edema no dorso dos pés e das mãos, hipoplasia do lado esquerdo do coração ou coarctação da aorta e baixo peso. As características faciais se tornam grosseiras, incluindo hipertelorismo ocular, ptose palpebral, pregas epicânticas, sobrancelhas espessas, além de uma baixa implantação de cabelos na nuca e tórax em barril. Ao longo do desenvolvimento, a baixa estatura se torna mais evidente e as pacientes com síndrome de Turner exibirão um atraso na maturação sexual, amenorreia primária ou secundária, disgenesia gonadal (gônadas em fita), mamas pouco desenvolvidas, aumento da distância entre os mamilos e escassez de pelos pubianos e axilares, além de malformações anatômicas. Em geral, as mulheres com a síndrome têm inteligência próxima do normal, mas pode-se observar em alguns casos um atraso no desenvolvimento intelectual e verbal somados a uma diminuição da capacidade de percepção espacial e dificuldades de socialização. Como mencionado anteriormente, mulheres com cariótipo 45,X, cujo cromossomo X é de origem paterna, costumam apresentar habilidades cognitivas sociais superiores em relação àquelas cujo cromossomo X é de origem materna. Nas mulheres adultas, a infertilidade é uma queixa constante, mas a gravidez em poucas portadoras da síndrome pode se tornar possível, provavelmente em decorrência da presença de mosaicismo gonadal. Nos casos de síndrome de Turner, nos quais ocorre deleção parcial do segundo cromossomo X, a fertilidade pode ser mantida, mas há risco de formação de gametas com a mesma alteração.

O tratamento das pacientes com síndrome de Turner visa à reposição de estrogênio a partir da puberdade, a fim de auxiliar no desenvolvimento das características sexuais secundárias. Além disso, a administração de hormônio do crescimento recombinante, quando implementada no momento correto, pode melhorar em aproximadamente 6 cm a altura final das pacientes.

Síndrome de Klinefelter (47,XXY e demais cariótipos atípicos)

A *síndrome de Klinefelter* é originada pela presença de uma ou mais cópias extras do cromossomo X na presença do cromossomo Y em um indivíduo do sexo masculino. A condição atinge cerca de 1 a cada 1.000 nativivos, mas acredita-se que essa frequência seja subestimada, em decorrência das consequências fenotípicas brandas a ela associadas. Em geral, a síndrome não é identificada clinicamente na infância, visto que os portadores em geral exibem os primeiros indícios de hipogonadismo na puberdade. Nessa fase, os testículos se mantêm pequenos e as características sexuais secundárias não se desenvolvem adequadamente. Os pacientes costumam ser altos e magros com membros longos e em alguns deles, observa-se um quadro de ginecomastia, o que leva os portadores da síndrome a terem maior propensão ao câncer de mama em relação aos homens 46,XY. A distribuição da gordura corporal exibe um padrão tipicamente feminino e os pelos corporais são escassos. Além disso, a compreensão verbal e o desempenho intelectual podem estar um pouco abaixo da média. O reconhecimento clínico da síndrome acontece quase sempre na idade adulta, em decorrência de queixas de infertilidade e tanto oligo como azoospermia podem estar presentes. Na maioria dos casos, o cariótipo associado à síndrome é o 47,XXY, e a não disjunção neste caso tem iguais chances de ter ocorrido durante a gametogênese materna ou paterna. Quando o cromossomo X extra é de origem paterna, em geral ocorreu uma falha na recombinação entre as regiões pseudoautossômicas dos cromossomos X e Y durante a meiose I. Quando o cromossomo X extra é de origem materna, a maioria dos casos é decorrente de falhas na meiose I materna, que são influenciadas pela idade materna avançada. Os demais erros acontecem durante a segregação na meiose II e no caso de mosaicismo, na mitose após a formação do embrião. Raramente podem ser observadas aneuploidias variantes da configuração 47,XXY clássica que incluem os cariótipos 48,XXXY, 48,XXYY e 49, XXXY. Mesmo que todos os cromossomos X adicionais sofram inativação, os cariótipos atípicos costumam estar relacionados a uma maior severidade, com maior comprometimento do desenvolvimento sexual e deficiência intelectual. Cerca de 15% dos pacientes com a síndrome exibem mosaicismo, sendo a combinação 46,XY/47,XXY a mais frequentemente encontrada. Além dos cariótipos citados, em raros casos, o fenótipo da síndrome de Klinefelter, com exceção da estatura elevada, já foi observado em homens com o cariótipo 46,XX, nos quais o gene SRY foi translocado para um dos cromossomos X, como uma consequência de uma recombinação meiótica anormal fora das regiões pseudoautossômicas.

O tratamento dos pacientes com a síndrome de Klinefelter visa à reposição de testosterona após a puberdade, que contribui para um melhor desenvolvimento das características sexuais secundárias. Além disso, embora o desenvolvimento dos testículos no portador da síndrome de Klinefelter pareça ser normal, existe perda significativa de células germinativas durante a maturação puberal. Sendo assim, a espermatogênese pode estar preservada em apenas uma pequena parcela de túbulos, os quais podem ser alvos de técnica de reprodução artificial. No entanto, aspectos éticos e legais devem ser avaliados previamente à utilização destas metodologias.

Síndrome do Triplo X (47,XXX e demais cariótipos atípicos)

A *trissomia do X* ou *síndrome do Triplo X* está associada à presença de uma ou mais cópias extras do cromossomo X, na ausência do cromossomo Y, em um indivíduo do sexo feminino. O cariótipo mais comum que está presente em 1 a cada 1.000 nativivas é a configuração 47,XXX, e é decorrente em 95% dos casos de não disjunção na meiose I materna, com forte influência da idade materna avançada. Em casos raros, cariótipos atípicos (48,XXXX e 49,XXXXX) também podem ser encontrados.

Embora exibam uma estatura um pouco acima da média e possam apresentar uma inteligência ligeiramente diminuída, em geral, as mulheres com cariótipo 47,XXX são fenotipicamente normais e passam desapercebidas na população. As mudanças puberais costumam acontecer na idade esperada e somente 20% das mulheres apresentam atraso na menarca, oligomenorreia, amenorreia secundária e atraso no desenvolvimento de características sexuais secundárias. Geralmente as mulheres são férteis, lembrando que há chance de gerar uma progênie cromossomicamente anormal. Contudo, alguns autores consideram que o cromossomo X extra pode ser segregado preferencialmente para o glóbulo polar I na meiose, aumentando as chances

das mulheres 47,XXX terem filhos normais. Nas portadoras de tetrassomia (48,XXXX) e pentassomia do X (49,XXXX) ocorre deficiência intelectual mais grave associada a demais manifestações físicas. Nestes casos, os cromossomos X adicionais são em geral de origem materna, como consequência de não disjunções consecutivas nas meioses I e II.

Síndrome do duplo Y (47,XYY)

A *síndrome do duplo Y*, historicamente conhecida como a síndrome do "*super macho*", surge da presença de uma ou mais cópias extras do cromossomo Y, representadas mais comumente pelo cariótipo 47,XYY, o qual está presente em um a cada 1.000 homens nascidos vivos. Os portadores da síndrome são fenotipicamente indistinguíveis dos homens 46,XY e geralmente são identificados em programas de triagem neonatal. A única característica associada à constituição 47,XYY é a estatura elevada e em alguns casos, podem existir problemas comportamentais, dificuldade de aprendizado, déficit de atenção, hiperatividade e impulsividade. No entanto, diferentemente do que pensou durante anos, quando a frequência da síndrome era avaliada em prisões e hospitais psiquiátricos, a síndrome não está associada à agressividade.

A não disjunção durante a meiose II da espermatogênese levando à formação de espermatozoides YY é a explicação mais plausível para o surgimento da síndrome do duplo Y. Curiosamente, os indivíduos com cariótipo 47,XYY não tem sua fertilidade alterada e não apresentam um risco aumentado de ter uma progênie cromossomicamente anormal. Isso é atribuído ao fato de o cromossomo Y extra ser eliminado na maioria das células que darão origem aos espermatócitos, possibilitando um pareamento normal dos cromossomos X e Y durante a meiose. Assim como visto para as demais aneuploidias sexuais, portadores de cariótipos mais raros (48,XYYY e 49,XYYY) apresentam fenótipos mais graves, que incluem deficiência intelectual, dismorfias faciais e anormalidades esqueléticas. Nestas variantes incomuns, a explicação mais comum para o surgimento da alteração seria uma não disjunção sequencial na meiose I e II durante a formação dos gametas masculinos.

Alterações no desenvolvimento gonadal

Durante o início do desenvolvimento embrionário, as gônadas em desenvolvimento independente do sexo genético (XX ou XY) são *bipotenciais*, ou seja, têm a habilidade de se transformarem em células masculinas (células de Sertoli e células Leydig) ou femininas (células foliculares e células Teca). A partir do 6º mês, as células germinativas primordiais migram para as saliências genitais, onde formarão o par de gônadas. É aceito que a feminilidade atribuída à formação do ovário é uma consequência natural do desenvolvimento gonadal, a menos que o *fator testículo-determinante* (TDF), codificado pelo gene determinante sexual do cromossomo Y (SRY) esteja presente. Na presença do cromossomo Y, o SRY iniciará a formação dos testículos pela ativação de fatores de transcrição específicos do sexo masculino que possibilitarão a proliferação e diferenciação das células bipotenciais. Caso o cromossomo Y esteja ausente, a gônada se diferenciará em ovário a partir da 8ª semana de gestação e a ovogônia começará a se desenvolver em folículos, entrando em meiose I por volta do terceiro mês de desenvolvimento intrauterino. Sendo assim, os portadores de aneuploidias sexuais, nas quais um cromossomo Y aparece em meio a duas ou mais cópias do cromossomo X serão sempre fenotipicamente do sexo masculino. No entanto, homens com cariótipo 46,XX podem ser observados, em decorrência da translocação do gene SRY para um ou ambos os seus cromossomos X. De maneira similar, portadores de cariótipos 46,XY ou 47,XXY com fenótipo feminino são resultantes da ausência ou mutação do gene SRY em seus cromossomos X. Essas alterações podem ocorrer em virtude de erros na recombinação que ocorre na meiose I. Nesta fase, é esperado que os cromossomos X e Y alinhem-se e troquem material genético entre suas regiões pseudoautossômicas. No entanto, em algumas situações a recombinação pode ocorrer fora destas regiões, originando indivíduos com sexo invertido (*homens XX e mulheres XY*). Contudo, a presença ou ausência do gene SRY não explica todos os casos de alterações na determinação gonadal e muitos outros fatores causais (alterações genéticas e não genéticas) ainda necessitam ser identificados.

Em alguns recém-nascidos, a caracterização sexual se torna difícil por causa da presença de *genitália ambígua*, condição a que se dá o nome de *hermafroditismo*. Nesses casos, podem

ser reconhecidas duas situações: a hermafroditismo verdadeiro e o pseudo-hermafroditismo. No *hermafroditismo verdadeiro*, existe tecido testicular e ovariano como órgãos separados ou fundidos em uma única gônada e, em geral, a genitália é ambígua. Em geral, a maioria dos hermafroditas verdadeiros é 46,XX e uma pequena parcela dos casos é atribuída a quimeras envolvendo linhagens celulares 46,XX e 46,XY. No *pseudo-hermafroditismo*, o tecido gonadal é apenas de um sexo e os órgãos genitais de outro. Os *pseudo-hermafroditas masculinos* (46,XY) têm genitália feminina ou incompletamente masculina, embora os testículos estejam alojados no canal inguinal ou nos grandes lábios. A vagina termina em fundo cego e os pseudo-hermafroditas masculinos não menstruam e são estéreis. As causas relacionadas com esta condição incluem alterações relacionadas com a biossíntese e metabolismo da testosterona, anormalidades de gonadotrofinas e alterações das células-alvo dos andrógenos. Uma condição bem estudada relacionada com o quadro de pseudo-hermafroditismo masculino é a síndrome de insensibilidade a androgênios. Nessa condição, os testículos secretam os androgênios normalmente, mas a ausência de receptores androgênicos nas células-alvo impede a diferenciação masculina completa.

Os *pseudo-hermafroditas femininos* são cariotipicamente e internamente mulheres (46,XX), mas a genitália externa é masculinizada ou ambígua, com hipertrofia do clitóris e fusão labial, que leva à formação de uma estrutura semelhante a bolsa escrotal. A principal causa para esta condição (50% dos casos) é a hiperplasia adrenal congênita, que ocorre em virtude da deficiência de enzimas do córtex adrenal fetal, envolvidas na biossíntese de cortisol (Tabela 4.6). O decréscimo da produção de cortisol provoca um aumento do hormônio adrenocortical (ACTH), causando aumento do córtex adrenal (hiperplasia adrenal) e estímulo na produção de esteroides adrenais pré-natais. Embora o desenvolvimento ovariano seja normal nestas mulheres, a produção de androgênios em excesso leva à virilização da genitália externa. Outra causa de pseudo-hermafroditismo feminino está relacionada com a presença de andrógenos na placenta ou pela ingestão de medicamentos pela mãe durante a gravidez, visando à prevenção de abortos.

■ **Tabela 4.6** Principais causas de pseudo-hermafroditismo feminino.

Virilização de causa fetal	• Hiperplasia adrenal congênita (deficiência de 21-hidroxilase; deficiência de 11β-hidroxilase; deficiência de 3β-hidroxisteroide desidrogenase) • Deficiência de aromatase • Resistência a glicocorticoide • Deficiência de P450 oxidoredutase
Virilização de causa materna	• Uso de drogas (esteroides anabolizantes, progestágenos, danazol) • Tumor de ovário ou de adrenal • Hiperplasia adrenal congênita virilizante materna não controlada
Síndromes dismórficas	–

Fonte: Castro, M.; Elias, L.L. Causas raras de pseudo-hermafroditismo feminino: quando suspeitar? *Arq Bras Endocrinol Metab* 2004; 49:126-137.

■ Diagnóstico pré-natal nas alterações cromossômicas e aconselhamento genético

O *diagnóstico pré-natal* representa um conjunto *multiprofissional* de procedimentos clínicos e laboratoriais que visam à detecção de anomalias genéticas ainda no desenvolvimento intrauterino. Nesse item, iremos nos ater ao *diagnóstico pré-natal* com enfoque citogenético. Durante os exames pré-natais de rotina, as gestantes são submetidas a triagens bioquímicas no soro materno e a testes de imagem com base em ultrassonografia para avaliar o desenvolvimento anatômico do feto. No entanto, em geral, existem três situações, nas quais a utilização de procedimentos invasivos de diagnóstico pré-natal é aconselhada. A primeira delas ocorre

quando existe *suspeita de anomalias citogenéticas* com base nos exames de rotina pré-natal. É então, solicitado que a gestante realize uma investigação genética mais detalhada. Além disso, alguns casais já têm *conhecimento prévio* de que a gestação envolve um risco aumentado de anomalia fetal, seja em virtude de um histórico familiar de cromossomopatia (ou de alteração balanceada), seja pela exposição da mãe a agentes mutagênicos. Para esses casais também é aconselhado o diagnóstico pré-natal. Por último, *gestantes com idade avançada* (superior a 35 anos) também estão sob risco de gerar embriões com aneuploidias, principalmente com trissomia do 21, e em alguns países, é praxe que essas mulheres sejam submetidas ao diagnóstico pré-natal. Nesse aspecto, devemos levar em consideração que, como a maioria das mulheres se reproduz antes desta idade e para elas os testes invasivos não são rotineiros, grande parte das crianças com aneuploidias autossômicas, como a síndrome de Down, somente é detectada no período pós-natal.

Durante o primeiro e o segundo trimestres de gestação são realizadas triagens de rotina no soro materno para marcadores proteicos específicos, que podem sugerir a presença de trissomias. No primeiro semestre (11ª a 13ª semana de gestação), são medidas as concentrações de proteína A plasmática associada à gestação (PAPP-A) e do hormônio gonadotrofina coriônica humana (hCG), tanto como hCH total como sua subunidade β livre. A PAPP-A costuma estar abaixo do limite normal em todas as trissomias, enquanto a hCG (ou β-hCG livre) mostra-se acima do limite na trissomia do 21, mas diminuída nas outras trissomias. Durante o segundo trimestre gestacional, são rastreados 3 marcadores séricos: a alfa feto proteína (MSAFP), a β-hCG livre e o estriol não conjugado (uE_3), que compõem a *triagem tripla*. Em alguns laboratórios, outro marcador, a inibina A, é adicionado à avaliação, que passa a ser uma *triagem quádrupla*. Os 4 marcadores estão diminuídos em todas as trissomias, com exceção da inibina A, a qual está aumentada na trissomia do 21 e permanece inalterada nas demais trissomias e, da β-hCG livre que está aumentada na trissomia do 21 e diminuída nas outras trissomias.

Entre a 11ª e a 14ª semanas de gestação, é realizado ainda o exame ultrassonográfico de *translucência nucal*. Nesse exame não invasivo, é medida a espessura do subcutâneo da nuca fetal em relação à idade gestacional. Essa medida está aumentada nas trissomias, em virtude da formação de edema nessa região da nuca, comum nas trissomias do 13, 18, 21 e monossomia do X. A capacidade do exame ultrassonográfico de identificar alterações requer a interpretação por profissionais experientes, além de equipamentos com resolução adequada. Atualmente, novos marcadores ultrassonográficos vêm sendo desenvolvidos para minimizar os resultados falso-negativos ou falso-positivos, como a medida do osso nasal, que está ausente em grande parte dos fetos com trissomia do 21 durante o primeiro trimestre de gestação. Somado a isso, a introdução da ultrassonografia obstétrica tridimensional e da ressonância nuclear magnética fetal tem possibilitado o aperfeiçoamento na visualização da anatomia fetal e na identificação de síndromes bem caracterizadas, embora ainda não sejam procedimentos de rotina pelo alto custo.

A combinação dos riscos obtidos pelas triagens bioquímicas do primeiro e segundo trimestres e dos exames ultrassonográficos pode sugerir a necessidade de realização de testes invasivos. Nesses testes, a coleta de material biológico para análise laboratorial é feita diretamente do material fetal.

Três testes invasivos são utilizados para o diagnóstico pré-natal de alterações citogenéticas: a *amniocentese*, a *punção de vilosidades coriônicas* e a *cordocentese*.

A escolha do método invasivo dependerá do período gestacional em que a suspeita de alteração genética foi levantada.

A *amniocentese*, realizada entre a 15ª e 16ª semanas de gestação, consiste na coleta de líquido do saco amniótico por punção transabdominal guiada por ultrassonografia. As células fetais flutuantes presentes no líquido amniótico são então cultivadas em laboratório para a avaliação de alterações cromossômicas ou gênicas. Também podem ser dosadas proteínas específicas como a alfafetoproteína (AFP). O risco de aborto oferecido pelo procedimento varia entre 0,5% a 1%, além de poder dar início a outras complicações como infecções, perda de líquido amniótico e lesão no feto causada pela agulha utilizada no procedimento.

A *amostragem de vilosidades coriônicas* pode ser realizada entre a 10ª e a12ª semana de gestação e envolve a retirada de fragmentos da área vilosa do córion que pode ser feita por via transabdominal ou transcervical. Em seguida, o material pode ser utilizado para testes bioquímicos, citogenéticos e moleculares. O risco de aborto gira em torno de 0,6 a 1,6%, em razão do risco de infecções, da ruptura de membranas e de outras causas, mas a punção de vilosidades coriônicas tem uma vantagem em relação à amniocentese de fornecer um resultado mais precoce. A técnica pode gerar diagnósticos duvidosos em decorrência de fatores que incluem a presença de mosaicismo somático confinado à placenta e casos de gemelaridade nos quais um dos fetos está vivo e outro em reabsorção.

Por último, a *cordocentese* representa a coleta de sangue fetal por punção do vaso umbilical. O procedimento, que pode ser realizado a partir da 18ª semana de gestação, é guiado por ultrassonografia e envolve a introdução de uma agulha na cavidade amniótica para retirada do material. O risco de aborto é de 2% a 5%.

Além dos testes acima descritos, nos últimos anos o advento de novas metodologias moleculares como a reação em cadeia da polimerase digital (PCR digital) e o sequenciamento de última geração têm aberto novos caminhos para a análise não invasiva de alterações cromossômicas fetais com base no DNA fetal circulante no sangue materno. A elaboração de novas técnicas não invasivas ganha mais relevância ainda, quando consideramos que houve um aumento significativo da idade materna na primeira gestação em países desenvolvidos, o que é acompanhado de maior risco de anomalias cromossômicas fetais. Corroborando esta observação, uma análise recente do *English and Welsh National Down Syndrome Cytogenetic Register* revelou um aumento de 71% no número de casos de síndrome de Down no período de 1989-2008, sem que houvesse aumento paralelo das taxas de nascimento.

Para famílias, nas quais uma alteração genética específica está segregando, o *diagnóstico genético pré-implantação* oferece alternativa às técnicas descritas, embora envolva questões éticas que ainda necessitam ser mais bem discutidas. Esse tipo de procedimento é um método de diagnóstico pré-natal precoce, no qual embriões originados por fertilização *in vitro* são avaliados quanto à presença de dada alteração genética e somente aqueles livres da alteração são transferidos para o útero. O diagnóstico é feito por técnicas de micromanipulação que envolvem a retirada de uma única célula do blastômero (6 a 8 células), seguida da análise molecular do material. Ainda não se sabe exatamente se a retirada dessas células pode ocasionar alguma consequência para o indivíduo no futuro, em virtude do reduzido número de casos de diagnóstico pré-implantação. Os embriões não utilizados são descartados como nos procedimentos de fertilização *in vitro* de rotina, o que provoca discussões éticas sobre a possibilidade desta seleção ser um método de aborto.

É válido ressaltar que a identificação de um cariótipo normal nas diferentes técnicas de diagnóstico pré-natal não é certeza de que o feto não apresente uma alteração genética. No diagnóstico pré-natal, em geral, são avaliadas alterações em nível citogenético ou mutações em genes específicos, para os quais havia um conhecimento prévio de histórico familiar. Por outro lado, nos casos em que for identificada alguma alteração genética, o ideal é que a alteração envolvida seja confirmada no período pós-natal.

A comunicação do resultado de um exame pré-natal é uma tarefa difícil e deve ser realizada por profissionais especializados.

O diagnóstico do feto/embrião é de extrema valia, pois possibilita que se tracem estratégias voltadas para o prognóstico e para possíveis opções terapêuticas. Em algumas situações envolvendo defeitos cardíacos congênitos, por exemplo, já é possível que o tratamento ou correção sejam iniciados antes do nascimento do bebê, caracterizando a *Medicina Fetal*. Além disso, o diagnóstico pré-natal é útil para estimar os riscos de reaparecimento da mesma condição em uma nova gestação ou em uma futura geração, o que dá-se o nome de risco de recorrência. No entanto, grande parte das alterações detectadas no período embrionário é de natureza esporádica e não apresenta risco de reincidir diferente daquele esperado para a população em geral.

Risco de recorrência
Chance de uma condição genética detectada em um ou mais membros da família aparecer em outro familiar na mesma progênie ou em gerações seguintes

O *aconselhamento genético* é definido como um processo que tem por objetivo final ajudar pessoas a entender e se adaptar às implicações médicas, fisiológicas e familiares da contribuição

genética para uma doença. O processo integra: (a) a interpretação das histórias familiares e médicas para acessar a chance de ocorrência ou recorrência da doença; (b) a educação sobre a herança, testes disponíveis, manejo, prevenção, recursos e pesquisa; (c) o aconselhamento para promover escolhas e adaptação ao risco ou à condição.

Com isso, o aconselhamento genético não somente se limita a fornecer informações a respeito de uma doença genética e de seus riscos de recorrência, mas também possibilita um acompanhamento do impacto psicológico na família. Além disso, o aconselhamento genético possibilita à família afetada por uma doença genética se planejar quanto ao nascimento de novos membros. Nesse processo não diretivo, não se deve dizer aos pacientes que decisão tomar (ter ou não filhos, esterilização, inseminação artificial, adoção); o dever do conselheiro genético é informar as opções disponíveis. Isso também é válido para as opções de tratamento, se disponíveis. Desse modo, ao contrário dos princípios eugênicos, o aconselhamento genético deve se concentrar nos interesses dos indivíduos afetados por uma doença genética e seus familiares, e não nos da sociedade. No passado, o aconselhamento genético era realizado exclusivamente por médicos. No entanto, a evolução do processo exigiu a necessidade da articulação entre vários profissionais da área de saúde.

Na maioria das vezes, o aconselhamento genético é realizado quando já existe um indivíduo afetado por uma doença genética na família, tendo caráter retrospectivo; o ideal seria que a família de alto risco fosse identificada antes do aparecimento de um membro afetado, de maneira prospectiva. Devemos ter em mente que, apesar de já existirem formas de se tratar doenças genéticas, como é o caso das aminoacidopatias (p. ex., fenilcetonúria), a maioria delas carece de tratamentos terapêuticos específicos e a melhor forma de evitá-las ainda é a prevenção. Enquanto ferramentas de prevenção ainda não estão disponíveis de modo global, a utilização de testes pré-natais ou perinatais efetivos no diagnóstico de doenças genéticas também tem grande importância para o tratamento precoce e o planejamento familiar.

Para concluir, neste capítulo procuramos abordar como os processos de divisão celular acontecem e como a perturbação desses processos, mesmo diante da atuação de diferentes pontos controladores, pode resultar em alterações cromossômicas numéricas e estruturais. O diagnóstico dessas alterações e o reconhecimento de como elas surgiram tem se tornado cada vez mais preciso graças à evolução de metodologias citogenéticas que possibilitam a identificação dessas alterações até mesmo no período pré-natal. A comunicação do diagnóstico de uma alteração cromossômica envolve um conjunto de procedimentos multidisciplinar, visando, acima de tudo, à qualidade de vida do paciente e de sua família. Nesse processo, o profissional de genética deve estar ciente das normas éticas inerentes a sua profissão, de forma a respeitar a autonomia do paciente e a privacidade da informação genética.

RESUMO

- Antes de se dividir, uma célula precisa duplicar o seu conteúdo genético. À alternância entre a fase de duplicação (intérfase) e a divisão celular (fase M) dá-se o nome de ciclo celular
- Durante o ciclo celular, para que um cromossomo se mantenha viável, ele precisa apresentar dois telômeros, uma origem de replicação, no mínimo, e um centrômero. A movimentação ordenada dos cromossomos durante a divisão celular se dá pela ligação dos cinetócoros de seus centrômeros aos microtúbulos do fuso
- As células humanas têm dois mecanismos de divisão celular: a mitose e a meiose. A mitose é utilizada pelas células somáticas diploides para o crescimento celular e a regeneração dos tecidos, ao passo que a meiose visa à formação de gametas haploides com base em células diploides da linhagem germinativa
- Na mitose, há distribuição igualitária do conteúdo genético da célula-mãe diploide vinda da intérfase para as duas células-filhas diploides, pela separação das cromátides-irmãs na anáfase. Em seguida, as células-filhas geradas entram em intérfase novamente
- Pontos de checagem na mitose monitoram a formação do fuso, como também controlam a ligação de suas fibras aos cinetócoros associados aos centrômeros, garantindo que todos os cromossomos sejam capturados
- A melhor fase para avaliação dos cromossomos humanos é a metáfase mitótica. Nela, os cromossomos humanos podem ser diferenciados pelo tamanho, pela forma e pelo padrão de bandas. A ordenação dos cromossomos de um indivíduo aos pares e em ordem decrescente de tamanho recebe o nome de *cariótipo*
- Na meiose, ocorrem duas etapas subsequentes de divisão celular sem que haja replicação do conteúdo genético entre elas. Na primeira etapa (meiose I ou reducional), ocorre separação dos cromossomos homólogos durante a anáfase I, enquanto na

segunda (meiose II ou equacional), o número haploide é mantido por meio da separação das cromátides na célula parental durante a anáfase II. O conjunto diploide é restaurado com a formação do zigoto que, em seguida, sofre uma série sucessiva de ciclos celulares para formação do embrião

- Três eventos geram variabilidade gamética na meiose: o *crossing over* entre cromátides não irmãs de cromossomos homólogos, a segregação aleatória dos cromossomos homólogos na anáfase reducional e a separação das cromátides na anáfase equacional. Durante o *crossing over*, quanto mais próximos dois genes estiverem um do outro, maior a possibilidade de eles serem transmitidos ligados para o mesmo gameta
- Ao longo do processo evolutivo, uma barreira para o *crossing over* foi estabelecida entre os cromossomos X e Y e a recombinação entre esses cromossomos aconteceu apenas nas regiões pseudoautossômicas. A ausência de *crossing over* ao longo desses cromossomos fez com que eles adquirissem características próprias
- A formação dos gametas masculinos é iniciada na puberdade e ocorre continuamente, durante toda a vida, enquanto todos os gametas femininos estão pré-formados ainda no desenvolvimento intrauterino, sendo mais suscetíveis a alterações cromossômicas com o avançar da idade materna. Além disso, a espermatogênese gera 4 células funcionais, ao passo que, na ovocitogênese, apenas uma das células geradas apresenta funcionalidade
- As alterações cromossômicas compreendem anomalias numéricas e anomalias estruturais
- As alterações cromossômicas numéricas podem levar a célula a ter um número de cromossomos que é um múltiplo exato (euploidias) ou não (aneuploidias) do conjunto haploide. As euploidias podem surgir por dispermia, endomitose ou pela fecundação por gametas diploides. O principal mecanismo envolvido nas aneuploidias é a não disjunção cromossômica durante as anáfases meióticas ou mitótica
- As alterações cromossômicas estruturais surgem de uma ou mais rupturas cromossômicas transversais, seguidas de uma reconstituição cromossômica anormal, podendo envolver um ou mais cromossomos. Essas alterações podem acontecer espontaneamente ou ser desencadeadas por agentes externos (infecções virais, drogas mutagênicas). Alterações estruturais incluem deleções, duplicações, inversões, translocações e inserções
- Rearranjos balanceados que aparentemente não envolvem a perda de segmentos cromossômicos podem resultar em configurações cromossômicas desbalanceadas na gametogênese, por causa de erros no pareamento dos cromossomos homólogos
- Em geral, alterações que envolvem os cromossomos sexuais são fenotipicamente mais brandas em relação àquelas que envolvem os autossomos, em virtude da compensação de dose atribuída à inativação do cromossomo X.

AUTOAVALIAÇÃO

4.1 Quais estruturas um cromossomo eucariótico precisa ter para ser funcional?

4.2 Quais são as principais diferenças entre a anáfase da mitose e as anáfases das meioses reducional e equacional?

4.3 Discuta os três eventos que ocorrem na meiose e que podem ocasionar variabilidade genética em nível gamético.

4.4 Conceitue haplótipo e diferencie distância genética de distância física.

4.5 Distinga a espermatogênese da ovocitogênese.

4.6 Descreva como uma triploidia e uma tetraploidia podem surgir.

4.7 Por que podemos dizer que a idade materna avançada é um fator de predisposição à ocorrência de alterações cromossômicas numéricas na progênie?

4.8 Descreva os parâmetros utilizados para distinção dos cromossomos humanos.

4.9 Conceitue mosaicismo e quimeirismo.

4.10 Como deleções e duplicações cromossômicas podem surgir?

4.11 Defina translocação robertsoniana e demonstre como esse rearranjo em sua forma balanceada pode levar à síndrome de Down na progênie.

4.12 Por que os cromossomos acêntricos são perdidos durante a divisão celular?

4.13 Explique por que as alterações nos cromossomos sexuais tendem a ser fenotipicamente mais brandas do que aquelas envolvendo autossomos.

BIBLIOGRAFIA

BAKER, M. Genomes in three dimensions. *Nature* 2011; 470:289-294.

BEIGUELMAN, B. "El consejo genético", pp. 141-152. V Congresso Latino-Americano de Genética, 1982.

BENJAMIN, L. *Genes IX.* Porto Alegre: Artmed, 2009.

BROTHMAN, A. R.; PERSONS, DL; SHAFFER, L.G. Nomenclature evolution: changes in the ISCN from the 2005 to the 2009 edition. *Cytogenet Genome Res.* 2009;127:1-4.

CASTRO, M.; ELIAS, L.L. Causas raras de pseudo-hermafroditismo feminino: quando suspeitar?. *Arq Bras Endocrinol Metab* 2004; 49:126-137.

ENSEMBL GENOME BROWSER. 2012. Disponível em: http://www.ensembl.org/

GARDINER, K.; HERAULT, Y.; LOTT, I.T. *et al.* Down syndrome: from understanding the neurobiology to therapy. *The Journal of Neuroscience* 2010; 30: 14943-14945.

HAHN, S.; LAPAIRE, O.; TERCANLI, S.; KOLLA, V.; HÖSLI, I. Determination of fetal chromosome aberrations from fetal DNA in maternal blood: has the challenge finally been met?. *Expert Rev Mol Med.* 2011;13:e16.

HOUK, C.P.; ROGOL, A.; LEE, P.A. Fertility in men with Klinefelter syndrome. *Pediatr Endocrinol Rev* 2010; Suppl 1: 182-6.

KNOBLICH, J. Asymmetric cell division: recent developments and their implications for tumour biology. *Nature reviews – Molecular Cell Biology* 2010; 11:849-860.

MALUF, S.W.; RIEGEL, M. *et al. Citogenética humana.* Porto Alegre: Artmed, 2011.

MORRIS, J.K.; ALBERMAN, E. Trends in Down's syndrome live births and antenatal diagnoses in England and Wales from 1989 to 2008: analysis of data from the National Down Syndrome Cytogenetic Register. *British Medical Journal* 2009; 339: b3794.

MOZDARANI, H.; MEYBODI, A.M.; KARIMI, H. Impact of pericentric inversion of Chromosome 9 [inv (9) (p11q12)] on infertility. *Indian J Hum Genet* [serial online] 2007 [cited 2011 Apr 19];13:26-9. Disponível em: http://www.ijhg.com/text.asp?2007/13/1/26/32031.

NATIONAL SOCIETY OF GENETIC COUNSELORS' DEFINITION TASK FORCE, RESTA, R.; BIESECKER, B.B.; BENNETT, R.L.; BLUM, S.; HAHN, S.E.; STRECKER, M.N.; WILLIAMS, J.L. A new definition of Genetic Counseling: National Society of Genetic Counselors' Task Force report. *J Genet Couns.* 2006;15:77-83.

NUSSBAUM, R.L.; MCINNES, RR & WILLARD, H. F. *Thompson & Thompson. Genética médica.* 7ª ed., Rio de Janeiro: Elsevier, 2007.

ROSS, M.T.; BENTLEY, D.R.; TYLER-SMITH, C. The sequences of the human sex chromosomes. *Curr Opin Genet Dev.* 2006;16:213-218.

SNUSTAD, D.P.; SIMMONS, M.J. Tradução Paulo A. Motta. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

YOUNG, S.S.; ESKENAZI, B.; MARCHETTI, F.M.; BLOCK, G.; WYROBEK. The association of folate, zinc and antioxidant intake with sperm aneuploidy in healthy non-smoking men. *Human Reproduction* 2008; 1-9.

5

Herança Monogênica

Objetivos de estudo

Compreender os padrões monogênicos de transmissão gênica e saber diferenciá-los
Ser capaz de elaborar heredogramas por meio da simbologia pertinente
Entender a importância da consulta ao OMIM
Perceber como a consanguinidade aumenta o risco de traços genéticos recessivos
Conhecer os fatores e mecanismos que podem influenciar a expressão fenotípica de um gene, dificultando a interpretação do padrão de herança em uma genealogia
Saber diferenciar penetrância de expressividade variável
Perceber a diferença entre heterogeneidade alélica e heterogeneidade de *locus*
Compreender como ocorre a transmissão de genes localizados no DNA mitocondrial

Conceitos-chave do capítulo

Alelo
Antecipação
Árvore genealógica
Autossomo
Azoospermia
Casamento consanguíneo
Caso-índice
Coeficiente de endocruzamento
Congênito
Consanguinidade
Cromossomo X
Cromossomo Y
Cromossomos sexuais
Diploide
Dissomia uniparental
Dominante
Expressividade variável
Fenocópia
Fenótipo
Gameta
Gene *SRY*
Genes deletérios
Genótipo
Hemizigoto
Herança autossômica dominante
Herança autossômica recessiva
Herança dominante ligada ao X
Herança holândrica
Herança influenciada pelo sexo
Herança ligada ao Y
Herança limitada ao sexo
Herança materna
Herança matrilinear
Herança mendeliana
Herança mitocondrial
Herança recessiva ligada ao X
Heredograma
Heterogeneidade alélica
Heterogeneidade de *locus*
Heterogeneidade genética
Heterozigoto
Heterozigoto assintomático
Heterozigoto composto
Homozigoto
Idade de manifestação variável
Imprinting genômico
Locus
Mosaicismo
Mosaicismo de células somáticas
Mosaicismo germinativo
Mosaicismo gonadal
Mutação de novo
Mutação nova
Online Mendelian Inheritance in Man (OMIM)
Penetrância
Penetrância completa
Penetrância incompleta
Penetrância reduzida
Pleiotropia
Polidactilia
Princípio da segregação
Princípio da distribuição independente
Probando
Propósito
Recessivo

Introdução

Alelo
Forma alternativa de um gene que ocupa um *locus* cromossômico específico

Diploide
Organismo com duas cópias de cada cromossomo

A herança monogênica ocorre quando os alelos de um único gene do genoma nuclear de um organismo diploide são transmitidos dos genitores à sua prole. A transmissão dos caracteres monogênicos, também denominados mendelianos, segue o modelo de segregação descrito pelo monge austríaco Gregor Mendel em meados do século XIX. Em seus estudos com ervilhas, Mendel observou certa regularidade no padrão de ocorrência, em gerações sucessivas, de determinadas características nas plantas, o que o levou a estabelecer alguns conceitos genéticos

relativos ao modo como as unidades hereditárias responsáveis pelos fenótipos eram transmitidas, apesar de ele nada saber sobre genes, cromossomos ou meiose.

Além do reconhecimento de que certos fenótipos eram condicionados por fatores genéticos específicos, dois princípios centrais foram deduzidos com base nos estudos de Mendel: o da *segregação* e o da *distribuição independente*.

O primeiro estabelece que, na reprodução sexuada, os fatores genéticos herdados de cada genitor separam-se na formação dos gametas, sendo que apenas um alelo de cada par será transmitido à prole pelo genitor. Na verdade, o *princípio da segregação* descreve o que ocorre com os cromossomos na meiose, ou seja, os cromossomos segregam, possibilitando que genes ligados a eles sejam transmitidos como entidades distintas de geração a geração.

O *princípio da distribuição independente* estabelece que genes localizados em *loci* distintos são transmitidos independentemente, ou seja, a transmissão do alelo em um *locus* não interfere na transmissão do outro.

Mendel observou também que, no nível do fenótipo, os efeitos de um alelo em determinado *locus* cromossômico podiam mascarar os efeitos do outro alelo no mesmo *locus*, levando-o a conceituar as características como sendo dominantes ou recessivas. Surgiam, assim, os conceitos de *dominância* e *recessividade*.

Em virtude da localização cromossômica do gene, os traços monogênicos podem ser classificados como autossômicos, se forem determinados por um gene situado em um autossomo, ou ligados ao sexo, quando o gene estiver localizado em um dos cromossomos sexuais humanos (X e Y).

Os traços podem ser *dominantes*, codominantes ou *recessivos*, dependendo do efeito que um alelo possa ter sobre o outro, ou seja, a expressão de um alelo em heterozigose pode ou não mascarar o efeito de outro alelo no fenótipo. É importante salientar que os genes, em si, não são dominantes, codominantes ou recessivos. Essas denominações referem-se aos fenótipos, que são o resultado da expressão gênica. Desse modo, traços dominantes são aqueles decorrentes

Fenótipo
Característica observada em um organismo, determinada pela interação de seu genótipo com o ambiente em que se expressa

Gameta
Célula da linhagem reprodutiva (espermatozoide ou ovócito). Apresenta número haploide de cromossomos (em humanos, 23 cromossomos)

***Locus* (no plural, *loci*)**
Posição ocupada por um gene em um cromossomo

Dominante
Qualquer característica que se expressa fenotipicamente em heterozigotos para o alelo determinante do traço

Recessivo
Qualquer característica que se manifesta fenotipicamente em homozigotos ou em heterozigotos compostos para o alelo determinante do traço

Autossomo
Qualquer cromossomo nuclear diferente dos cromossomos sexuais X e Y. Existem 22 pares de autossomos em humanos

Codominante (codominância)
Quando os efeitos de dois alelos distintos presentes em um *locus* no par de cromossomos homólogos são observados no fenótipo

Heterozigoto (heterozigose)
Organismo com dois alelos diferentes em determinado *locus* no par de cromossomos homólogos

OMIM
Catálogo *on-line* de traços humanos com herança monogênica

G·F Genética em Foco

OMIM: Uma importante ferramenta na Genética Humana

Muitas doenças genéticas são decorrentes de mutações que ocorrem em um único gene, devendo ser ressaltado que grande parte do progresso alcançado na prática médica deve-se ao mapeamento, à clonagem e à análise molecular desses genes, que possibilitaram o conhecimento dos mecanismos fisiopatológicos associados a inúmeros distúrbios. O OMIM descreve mais de 19 mil genes e fenótipos monogênicos em humanos. Esse banco de dados é uma versão *on-line* do catálogo Mendelian Inheritance in Man (MIM), publicado pela primeira vez em 1966 por Victor McKusick, que vem sendo continuamente atualizado e reúne extensa informação sobre genes humanos e traços monogênicos, patológicos ou não. Por sua confiabilidade e praticidade, o OMIM é um referencial na área da genética humana. Cada entrada registrada nesse banco de dados é identificada por um número único composto por 6 dígitos, o número MIM, sendo que o primeiro dígito normalmente indica o modo de herança: 1 (*loci* ou fenótipos autossômicos dominantes); 2 (*loci* ou fenótipos autossômicos recessivos); 3 (*loci* ou fenótipos ligados ao X); 4 (*loci* ou fenótipos ligados ao Y); 5 (*loci* ou fenótipos mitocondriais); 6 (*loci* ou fenótipos autossômicos acrescentados após 1994). Outros símbolos podem aparecer antes do número, como [*], o qual indica que a sequência do gene em questão é conhecida. Tome-se como exemplo, a síndrome do X frágil, principal causa hereditária de deficiência mental ligada ao cromossomo X, identificada pelo número MIM 300624, que é decorrente de mutação no gene *FMR1*, cujo registro é MIM *309550.

Homozigoto (homozigose) Organismo com dois alelos iguais em determinado *locus* no par de cromossomos homólogos

Hemizigoto (hemizigose) Organismo com somente um alelo em determinado *locus* nas células diploides. Os homens são hemizigotos para genes ligados ao X por terem apenas um cromossomo X

da expressão de determinado gene em heterozigose, ou seja, a presença de um único alelo condiciona a característica. Fenótipos recessivos, em geral, são decorrentes de mutações que comprometem ambas as cópias funcionais do gene, levando à perda total ou parcial de sua função. Sendo assim, indivíduos portadores de traços recessivos são homozigotos para o gene em questão, excetuando-se os homens com traços recessivos ligados ao X, que são hemizigotos. Além dos traços dominantes e recessivos, alguns fenótipos, denominados codominantes, são consequência da presença de alelos que são igualmente expressos no estado heterozigoto, ou seja, a expressão de um não mascara a expressão do outro. É esse o caso do *locus* do grupo sanguíneo ABO, no qual as combinações possíveis entre três alelos (I^A, I^B e i) determinam a presença dos antígenos A e B nas hemácias. Os alelos I^A e I^B são codominantes entre si e dominantes em relação ao alelo i, que não codifica antígeno algum.

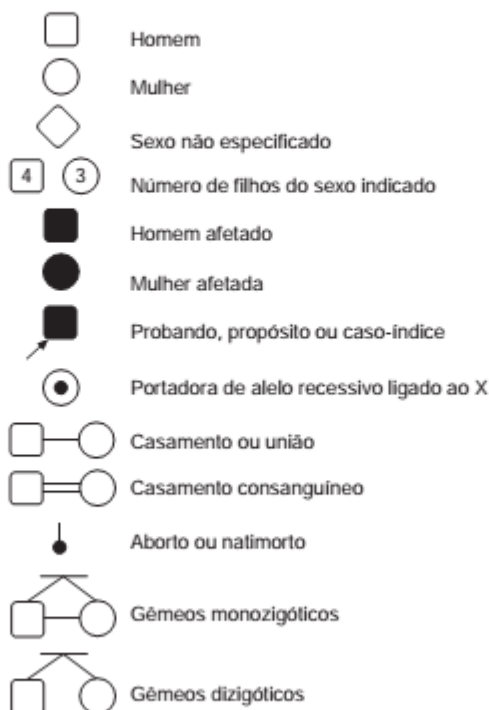
Heredogramas

Heredograma Diagrama usado para mostrar a transmissão familiar de um traço hereditário

A genética clássica analisa as relações entre seres humanos com base em seu parentesco e, dessa maneira, o padrão de transmissão de determinada característica monogênica pode ser reconhecido por meio do levantamento da história familiar do indivíduo portador daquele traço. Com a elaboração de um heredograma ou *árvore genealógica*, uma das ferramentas mais utilizadas na genética médica, que nada mais é do que a representação gráfica de uma história familiar, é possível acompanhar a transmissão de um traço de pais para filhos em múltiplas gerações de uma genealogia.

Desse modo, o levantamento das informações relativas à determinada condição que segrega em uma família constitui o primeiro passo para avaliar se a característica em questão é transmitida de modo monogênico. Da análise do heredograma, que deve conter o maior número possível de dados de indivíduos em diferentes gerações, será possível estabelecer o padrão de herança do traço. O pequeno tamanho das famílias, bem como a falta de informação relativa aos parentes mais distantes, é fator que pode dificultar ou mesmo impossibilitar a avaliação do padrão de transmissão gênica.

Probando (propósito ou caso-índice) Indivíduo que apresenta determinado traço genético e pelo qual a família começa a ser averiguada

Como elaborar uma árvore genealógica? A construção de um heredograma segue convenções internacionais e se dá pela utilização de símbolos específicos, como se pode verificar na Figura 5.1. Denomina-se probando, *propósito* ou *caso-índice* o indivíduo com base no qual

Representação da numeração dos indivíduos nas gerações de uma família. O probando (indivíduo II-3) é a terceira pessoa da geração II e tem 25 anos de idade.

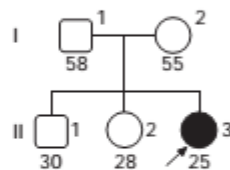



Figura 5.1 Símbolos comumente utilizados na elaboração de heredogramas.

se constrói a genealogia, sendo este indicado por uma seta. Desse modo, a elaboração de um heredograma começa com a obtenção de informações do probando e de seus parentes imediatos (pais e irmãos), sendo ampliado para gerações anteriores e posteriores. As gerações são representadas por algarismos romanos e os indivíduos em uma mesma geração são numerados da esquerda para a direita, em algarismos arábicos. A idade das pessoas pode ser registrada abaixo dos símbolos, e outras informações julgadas pertinentes devem ser anotadas.

■ Padrões de herança

Herança autossômica dominante

Esse padrão de herança é observado em muitos fenótipos mendelianos conhecidos, como acondroplasia, braquidactilia, polidactilia, doença do rim policístico do adulto, síndrome de Marfan, doença de Huntington, neurofibromatose tipo 1, distrofia miotônica, síndrome de Noonan e osteogênese imperfeita. Estima-se que a incidência de *distúrbios autossômicos dominantes* seja em torno de 1/1.000 nascimentos. Embora fenótipos dominantes sejam expressos em indivíduos heterozigotos ou homozigotos para o alelo mutante, em geral os portadores de traços dominantes são heterozigotos, resultantes da união de um genitor normal com um heterozigoto afetado, como exemplificado na Figura 5.2.

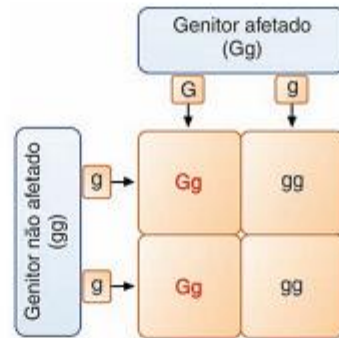



Figura 5.2 Quadrado de Punnett com as possíveis combinações alélicas que originam um fenótipo autossômico dominante.

Observe, na Figura 5.2, que, em média, metade da prole será heterozigota (Gg) e expressará o traço, ao passo que a outra metade será homozigota normal (gg). Os indivíduos homozigotos para o alelo mutante (GG), originados das raras uniões entre heterozigotos, em geral, desenvolvem fenótipos muito graves, associados a uma baixa expectativa de vida. É esse o caso da *acondroplasia* (nanismo com deformidades esqueléticas) que, quando em homozigose, é letal na infância.

Um heredograma típico da herança autossômica dominante pode ser encontrado na Figura 5.3.

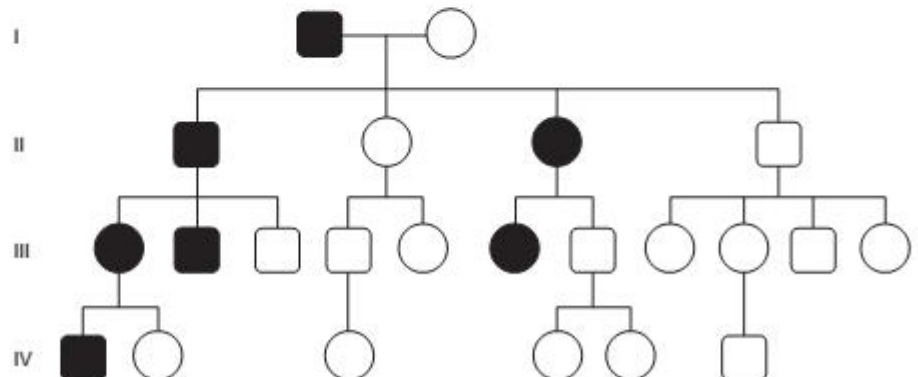



Figura 5.3 Padrão de herança autossômica dominante.

Observe que a doença é transmitida de uma geração para a outra de maneira contínua, ou seja, *não há saltos de gerações*. Por isso, diz-se que a herança dominante apresenta um padrão de *transmissão vertical*.

Como pode-se observar, um indivíduo com o distúrbio tem, geralmente, pai ou mãe afetada. Outra característica dos traços autossômicos é o fato de tanto homens como mulheres poderem igualmente apresentar a condição e transmiti-la a seus filhos de ambos os sexos. Qualquer pessoa afetada tem 50% de risco de transmitir o gene mutante para sua prole. Filhos ou irmãos saudáveis de indivíduos afetados não transmitem a doença a seus descendentes.

Herança autossômica recessiva

Um traço autossômico recessivo pode ocorrer igualmente em ambos os sexos, sendo que ele só se manifesta fenotipicamente quando o gene responsável estiver presente em homozigose. Assim, o indivíduo portador de um fenótipo recessivo recebe um gene mutante de cada genitor. Como mostrado na Figura 5.4, em geral, o pai e a mãe do probando são *heterozigotos assintomáticos* e apresentam, a cada gestação, 25% de chance de transmitir o caráter a sua prole. Como exemplos de doenças autossômicas recessivas, podemos citar: fenilcetonúria, retinite pigmentosa, síndrome de Hurler, síndrome de Bloom, albinismo, anemia falciforme, anemia de Fanconi, galactosemia, distrofia muscular do tipo cinturas, fibrose cística, alcaptonúria, doença de Gaucher e doença de Tay-Sachs.

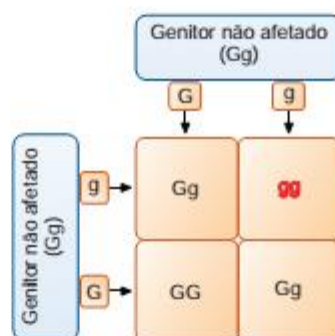



Figura 5.4 Quadrado de Punnett com as possíveis combinações alélicas que originam um fenótipo autossômico recessivo.

Como se pode verificar na Figura 5.5, o padrão de transmissão de distúrbios com esse tipo de herança genética é bem diferente daquele observado nos traços autossômicos dominantes.

É comum identificar, nas condições recessivas, a ocorrência da doença em um ou mais irmãos do probando, mas não em gerações passadas, ou seja, a doença, em geral, não acomete antepassados ou colaterais de indivíduos afetados.

Na verdade, a maioria dos casos de doenças autossômicas recessivas costuma ocorrer de maneira esporádica, e isso se deve, em parte, ao reduzido tamanho das famílias, em média com menos de quatro filhos por casal. Os irmãos normais do propósito apresentam risco de 2/3 de serem heterozigotos para o traço.

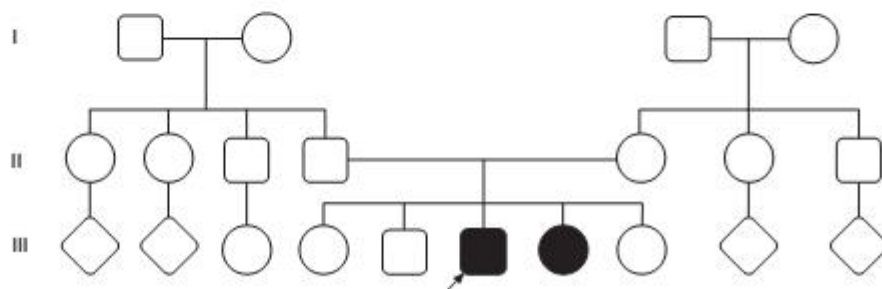



Figura 5.5 Padrão de herança autossômica recessiva.

Na Figura 5.5, pode-se constatar que os pais dos afetados não apresentam a doença, embora ambos sejam portadores do mesmo alelo mutante. Para você, isso é comum entre casais na população?

Na verdade, as chances de os genitores compartilharem um mesmo alelo mutante aumentam na proporção do *grau de parentesco* existente entre eles. Quanto mais rara for uma doença autossômica recessiva, maior a probabilidade de o indivíduo acometido ter pais *consanguíneos*. Dizer que dois indivíduos são consanguíneos significa inferir que eles, por apresentarem um ou mais ancestrais em comum, são geneticamente relacionados, o que aumenta a probabilidade de portarem mutações idênticas, herdadas de um mesmo antepassado (Figura 5.6).

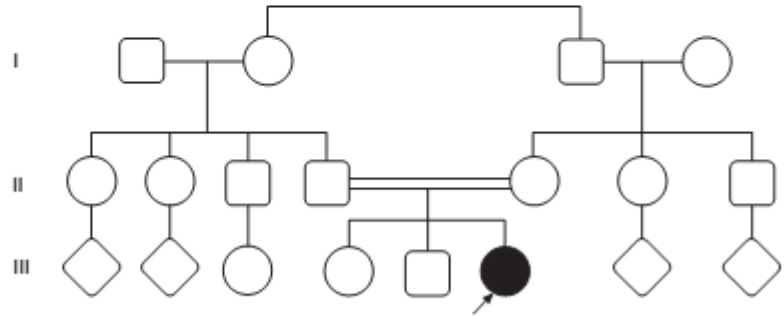



Figura 5.6 Padrão de herança autossômica recessiva com consanguinidade.

Podemos expressar o grau de consanguinidade pelo coeficiente de endocruzamento (F) definido como a probabilidade de que um indivíduo receba, em determinado *locus*, dois alelos idênticos por herança. Assim, o valor de F é proporcional ao grau de parentesco dos genitores e representa informação de grande importância na prática da genética médica, ao considerar que casamentos consanguíneos apresentam chance acima da média de produzir descendentes homozigotos para determinados genes deletérios.

Consanguinidade
Relação de parentesco

Coeficiente de endocruzamento
Probabilidade de uma pessoa ter dois alelos iguais por herança

Casamento consanguíneo
União entre pessoas aparentadas, ou seja, que têm um ou mais ancestrais em comum

Genes deletérios
Genes mutados que não desempenham adequadamente suas funções na célula, causando distúrbios ao organismo

Herança recessiva ligada ao X

O sequenciamento do cromossomo X, concluído no ano de 2005, revelou a presença de *1.098 genes codificadores de proteínas* e outros *150 genes codificadores de RNA*, distribuídos por, aproximadamente, 155 Mb, dos quais 54 genes têm homólogos funcionais no Y. O cromossomo X concentra 4% dos genes humanos e, apesar de seu pequeno conteúdo gênico em relação aos outros cromossomos de tamanho semelhante ao seu, ele ocupa lugar de destaque, considerando-se que cerca de 10% das doenças mendelianas conhecidas são causadas por mutações em genes ligados ao X. Esse cromossomo acumula um grande número de genes relacionados à função cerebral, o que explicaria a maior incidência de doenças neurocognitivas no sexo masculino em comparação ao feminino. Entre os genes que mapeiam no X, inclui-se o maior gene humano, com 2,22 Mb, codificador da proteína distrofina, relacionada às distrofias musculares de Duchenne e de Becker.

As características condicionadas pelos genes ligados ao X comportam-se de maneira diferenciada nos sexos masculino e feminino. A ausência quase total de homologia entre os cromossomos humanos X e Y possibilita afirmarmos que os genes localizados no X apresentam padrões de herança diferentes nos dois sexos em virtude de a mulher apresentar dois cromossomos X e o homem, apenas um. De modo semelhante à herança autossômica, os traços ligados ao X no sexo feminino mantêm as relações de dominância e recessividade, ao passo que, no sexo masculino, eles se expressam em *hemizigose*.

Como se pode verificar com a análise da Figura 5.7, nos traços recessivos ligados ao X, o sexo masculino é comumente mais afetado que o sexo feminino. Observe que não existe transmissão do fenótipo de um homem afetado para seus filhos homens, mas, sim, dele para todas as suas filhas, que serão portadoras. Elas terão 50% de probabilidade de transmitir o fe-

nótipo para seus filhos e igual probabilidade de transmitir o alelo deletério para suas filhas, que também serão portadoras. Entre as doenças mendelianas recessivas ligadas ao X, destacamos: hemofilias A e B, distrofia muscular de Duchenne, distrofia muscular de Becker, daltonismo, doença de Fabry, deficiência de G6PD, síndrome de Hunter, síndrome de Lesh-Nyhan e síndrome de Kennedy.

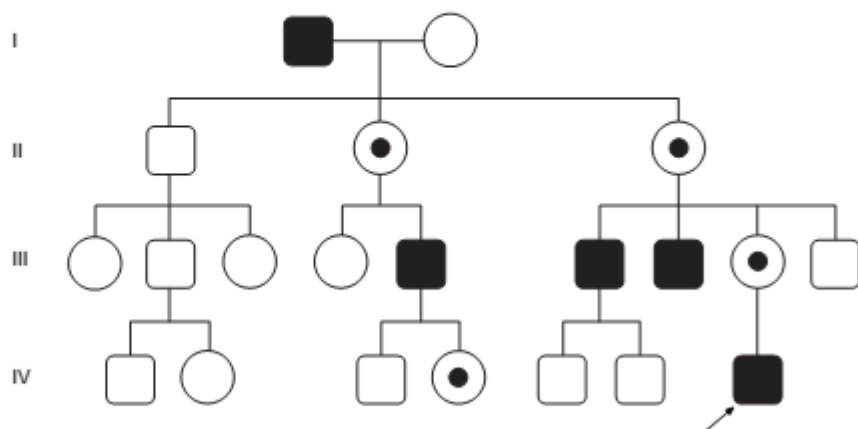



Figura 5.7 Padrão de herança recessiva ligada ao X.

Herança dominante ligada ao X

Enquanto os fenótipos recessivos ligados ao X em geral só ocorram em homens, os raros traços dominantes ligados a esse cromossomo são mais frequentemente encontrados nas mulheres, em virtude de elas poderem herdar o gene mutante do cromossomo X materno ou paterno. A principal característica observada nesse tipo de herança é em relação ao homem afetado transmitir o gene e o caráter para todas as suas filhas e para nenhum de seus filhos. As mulheres afetadas transmitem o traço para metade de seus filhos de ambos os sexos (Figura 5.8). Como exemplo, temos o raquitismo hipofosfatêmico, a síndrome de Rett, a síndrome de Goltz e a incontinência pigmentar. Cabe salientar que algumas doenças dominantes ligadas ao X cursam com fenótipos extremamente graves. É o caso da síndrome de Rett, que se manifesta como um quadro neurodegenerativo no primeiro ano de vida e acomete quase exclusivamente o sexo feminino, por ser letal nos homens hemizigotos. Existem alguns relatos da doença em indivíduos do sexo masculino com cariótipo 47,XXY, ou seja, portadores da síndrome de Klinefelter.

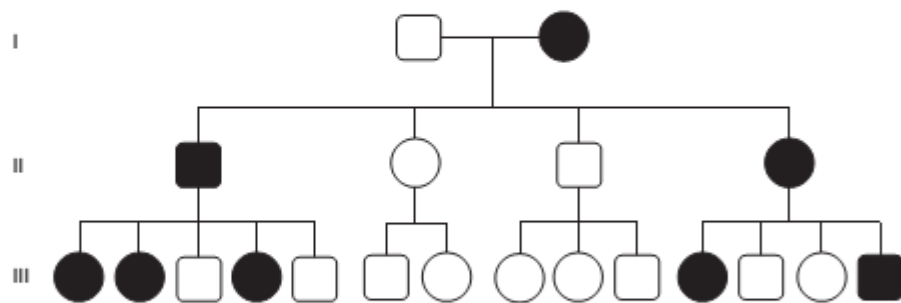



Figura 5.8 Padrão de herança dominante ligada ao X.

Herança holândrica
Padrão de herança de genes ligados ao cromossomo Y

Herança ligada ao Y (ou herança holândrica)

O cromossomo Y humano tem 60 Mb, o que representa cerca de 1% do conteúdo de DNA diploide. Pouco se sabia sobre a constituição gênica do Y até que o sequenciamento desse cromossomo, no ano de 2003, revelou ser ele formado por um mosaico de sequências *hete-*

rocromáticas e *eucromáticas* complexas e que, diferentemente do que se acreditava, contém *78 genes codificadores de proteínas*. Além do gene *SRY*, determinante do sexo masculino, os genes *DFNY1* (associado a um tipo de surdez) e *HEY* (a pelos nas orelhas), e diversos outros envolvidos na espermatogênese, mapeiam no Y. Estes, quando deletados, levam à infertilidade masculina em virtude da azoospermia.

Azoospermia
Ausência total de espermatozoides no sêmen

Como seria de se esperar, traços ligados ao Y têm sua transmissão exclusiva de homem para homem.

Herança limitada ou influenciada pelo sexo

Como mencionado, em relação aos traços autossômicos, recessivos ou dominantes, é de se esperar que homens e mulheres sejam igualmente afetados. Entretanto, alguns fenótipos autossômicos apresentam manifestação restrita em determinado sexo e existem traços que, embora se expressem em ambos os sexos, são mais frequentes em um deles, ou seja, sofrem a *influência do sexo*. Dois fenótipos dominantes bem conhecidos, nos quais verificamos manifestação preferencial em função do sexo são a *calvície pré-senil* e a *gota*, que afetam predominantemente os homens, provavelmente em virtude dos hormônios masculinos. Outro traço autossômico que sofre a influência do sexo é a *hemocromatose*, um distúrbio no metabolismo do ferro, de herança recessiva, que resulta no acúmulo desse mineral em diversos órgãos do corpo, o qual acomete homens na proporção de 5:1 em relação às mulheres. Acredita-se que, no sexo feminino, a perda de ferro pela menstruação contribua para a menor incidência do distúrbio entre as mulheres.

Como exemplos de *traços autossômicos limitados a determinado sexo*, podemos citar: esterilidade masculina decorrente de anormalidades testiculares (dominante) ou de hipogonadismo (recessivo); puberdade precoce no sexo masculino (dominante) e hiperplasia adrenal congênita (recessiva), uma importante causa de genitália ambígua em neonatos femininos.

As principais características que definem os padrões de transmissão monogênica estão resumidas na Tabela 5.1.

■ **Tabela 5.1** Características dos diferentes padrões de herança monogênica.

Herança	Características
I. Autossômica dominante	Transmissão vertical do fenótipo, sem pular gerações
	O indivíduo afetado tem um dos genitores afetados
	Homens e mulheres podem igualmente apresentar e transmitir o traço
	Afetados transmitem a condição para metade de sua prole
	Indivíduos normais não transmitem o traço
II. Autossômica recessiva	Transmissão horizontal do fenótipo, ou seja, presença do traço entre irmãos, mas não em gerações anteriores
	Pais dos indivíduos afetados, em geral, são heterozigotos e transmitirão o traço para 1/4 de sua prole
	Homens e mulheres têm igual probabilidade de apresentar e transmitir o traço
	Elevada taxa de consanguinidade entre os genitores
III. Recessiva ligada ao X	Geralmente apenas o sexo masculino é afetado
	O traço não é transmitido diretamente de homem para homem
	A transmissão, em geral, se dá por mulheres heterozigotas normais, e os homens afetados transmitem o gene mutante para todas as suas filhas, as quais serão portadoras
IV. Dominante ligada ao X	Os homens afetados transmitem o traço para todas as suas filhas e para nenhum de seus filhos, ou seja, não ocorre transmissão de homem para homem
	As mulheres afetadas heterozigotas transmitem a condição para metade dos filhos de ambos os sexos
	Homens e mulheres podem apresentar o traço, sendo este mais prevalente nas mulheres. Estas, por serem geralmente heterozigotas, apresentam fenótipo mais variável e brando em relação aos homens
V. Ligada ao Y	Somente homens são afetados
	Transmissão exclusiva de homem para homem

Variações nos fenótipos

Genótipo
Constituição genética de uma célula ou de um organismo. Mais especificamente, o termo é usado em referência aos alelos presentes em um *locus* cromossômico particular

Comumente, diz-se que toda regra tem exceções; no caso dos padrões de herança monogênica, não seria diferente. Na realidade, os crescentes conhecimentos de biologia molecular nos mostram claramente que existem muitos fatores e mecanismos que podem influenciar a expressão fenotípica de um gene. Sabemos que do genótipo ao *fenótipo* existe um longo caminho, no qual inúmeros passos metabólicos estão envolvidos, e que mesmo os fatores genéticos que contribuem para fenótipos monogênicos interagem com o genoma como um todo e, portanto, podem ter sua expressão fenotípica influenciada pela ação de outros genes, bem como de fatores ambientais externos e do próprio processo de envelhecimento.

Desse modo, é preciso conhecer e estar atento às variações que podem ocorrer na expressão de determinados fenótipos e que dificultam não apenas o diagnóstico daquela condição, mas também a interpretação das informações contidas no heredograma. A seguir, destacam-se alguns fatores que interferem nas correlações genótipo-fenótipo e que devem ser considerados na avaliação da transmissão de um traço monogênico.

Mutação nova e mosaicismo

Uma pessoa acometida por distúrbio autossômico dominante, em geral tem um dos genitores com o traço e outros familiares afetados. Entretanto, é possível que um indivíduo apresente determinado distúrbio genético, sem registro prévio de casos na família. Tal acontecimento pode ser causado por uma *mutação nova,* e isso é especialmente aparente em traços dominantes graves, nos quais, em geral, os genitores não são acometidos e não há história familiar do distúrbio.

Novas mutações, embora raras, surgem continuamente e ocorrem tanto em células somáticas quanto nas germinativas, levando ao aparecimento de doenças monogênicas em probandos cujos pais não são afetados.

Em alguns distúrbios dominantes, uma proporção significativa de indivíduos acometidos resulta de mutações novas e, como exemplo, temos a acondroplasia, causa mais comum de nanismo. Duas mutações *missense* na posição 1138 do gene do receptor 3 do fator de crescimento do fibroblasto (*FGFR3*) são responsáveis por mais de 80% dos casos da doença, sendo estes, filhos de genitores com estatura normal. Mutações novas no gene da distrofina também são causa frequente de distrofia muscular de Duchenne e novas mutações no gene *NF1* respondem por cerca de 50% dos casos de neurofibromatose tipo 1.

Mosaicismo
Refere-se à presença em um organismo de duas ou mais linhagens ou tipos celulares distintos derivados de um zigoto único

Além das novas mutações surgidas espontaneamente em um gameta paterno ou materno, a presença de mosaicismo nas células germinativas de genitores normais também contribui significativamente para a ocorrência de doenças em seus filhos. O mosaicismo para uma mutação específica surge na mitose durante a embriogênese e resulta na presença de dois ou mais tipos de células ou linhagem celulares, geneticamente distintas, porém oriundas de um mesmo zigoto, podendo afetar células somáticas ou germinativas. O *mosaicismo de células somáticas* pode ocasionar variações fenotípicas, estando, geralmente, associado a fenótipos mais brandos. O *mosaicismo germinativo ou gonadal,* por sua vez, resultante do surgimento de mutação em uma das células que dará origem à linhagem germinativa, responsável pela formação dos gametas, propicia o aparecimento de um distúrbio genético *de novo* na prole de indivíduos normais.

O fato de o indivíduo ter em suas gônadas células normais e células mutadas pode ocasionar a transmissão da mutação para seus descendentes. Embora o mosaicismo germinativo seja relativamente raro, ele aumenta significativamente o risco de recorrência do distúrbio na prole do indivíduo, a despeito de ele mesmo não ser afetado pelo traço, o que não ocorre com as mutações gaméticas esporádicas.

Uma forte evidência para suspeitarmos de que esse mecanismo esteja atuando ocorre quando dois ou mais filhos de um casal apresentam doença dominante ou ligada ao X, sendo seus pais normais e não existindo história familiar da doença. Um percentual significativo de casos de osteogênese imperfeita, acondroplasia, neurofibromatose tipo 1, hemofilia A e distrofia muscular de Duchenne resultam de mosaicismo gonadal.

Penetrância

Como vimos, por definição, indivíduos heterozigotos para mutações gênicas responsáveis por distúrbios dominantes manifestam o traço. Entretanto, uma pessoa pode ser heterozigota para determinada mutação dominante e esta não se manifestar no fenótipo, ou seja, um mesmo alelo pode se expressar fenotipicamente em uma pessoa e não produzir efeito detectável em outra. A essa característica denominamos penetrância, e ela, aliada à redução na prole de casais, representa um fator complicador na interpretação do heredograma, bem como na determinação dos riscos genéticos.

Penetrância
Refere-se à probabilidade que um gene apresenta de ter alguma expressão fenotípica

Define-se penetrância como sendo a probabilidade de o efeito de um gene se manifestar ou não ao nível de fenótipo. O valor da penetrância de um traço específico, estimado com base na *análise da segregação* desse traço em um grande número de famílias, expressa a relação entre os indivíduos que apresentam o gene para determinado fenótipo e aqueles que realmente exibem o caráter.

Traços que são expressos em 100% dos indivíduos portadores da mutação têm *penetrância completa* e aqueles, cuja expressão fenotípica ocorre apenas em parte dos indivíduos heterozigotos, apresentam *penetrância reduzida ou incompleta*.

Assim, se em uma população, 70% das pessoas portadoras de determinado gene mutante expressarem o fenótipo, a penetrância desse traço será de 70%. A penetrância incompleta é mais facilmente observada em traços autossômicos dominantes.

A Figura 5.9 ilustra a segregação de uma mutação autossômica dominante (representada por G) que causa a polidactilia. Observe que o fenótipo não se manifestou no indivíduo II-4, apesar de ele ser heterozigoto para o alelo mutante e ter transmitido o traço para dois de seus filhos (III-8 e III-9).

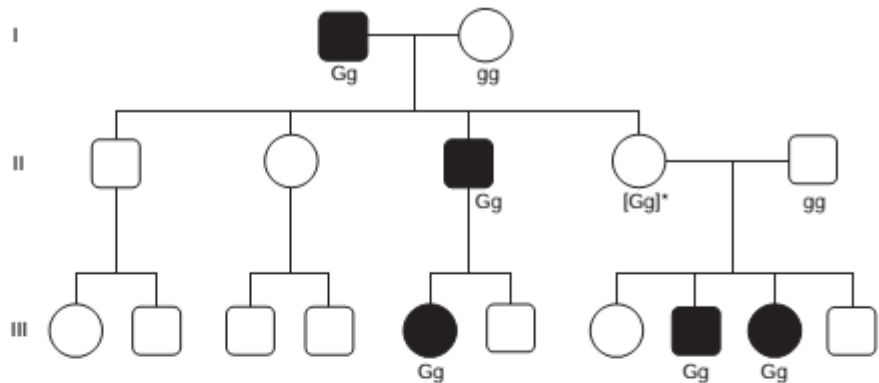



Figura 5.9 Segregação de fenótipo autossômico dominante com penetrância incompleta (polidactilia) G – alelo mutado relacionado com a polidactilia. [Gg]* – deduzimos tratar-se de um heterozigoto para o alelo G

Doenças genéticas de início tardio apresentam *penetrância relacionada com a idade*. Cabe aqui salientar que os traços genéticos não são, necessariamente, congênitos. Os fenótipos podem se manifestar em qualquer fase da existência, seja durante o desenvolvimento intrauterino, seja no nascimento ou no decorrer da vida. Todos conhecemos distúrbios que cursam com *idade de manifestação variável*, surgindo, muitas vezes, quando o portador do traço já procriou. Um exemplo clássico é a doença de Huntington, uma disfunção neurodegenerativa progressiva, extremamente grave, de herança autossômica dominante.

Congênito
Presente ao nascimento

Expressividade variável

Outro fator que deve ser considerado na avaliação de um fenótipo que está segregando em uma família refere-se à variação na sua expressão. Os mesmos fatores biológicos que influenciam a penetrância de um traço também podem contribuir para a manifestação fenotípica de modo diferenciado, até mesmo entre indivíduos de uma mesma família. É comum, em traços

autossômicos dominantes, observarmos acentuada variação nos efeitos fenotípicos produzidos por um mesmo gene em diferentes indivíduos, ou seja, uma mutação pode não resultar nas mesmas características clínicas em todos os seus portadores.

Expressividade variável
Variabilidade observada nos efeitos fenotípicos produzidos por um mesmo gene em diferentes indivíduos

A doença do rim policístico do adulto é um exemplo de fenótipo que apresenta expressividade variável, uma vez que algumas pessoas com a doença desenvolvem insuficiência renal no início da vida adulta, ao passo que outras têm somente alguns cistos nos rins que não afetam significativamente o funcionamento desses órgãos. Outros exemplos de doenças com expressão fenotípica variável são a neurofibromatose tipo 1, a osteogênese imperfeita, a fibrose cística e a síndrome de Waardenburg.

Antecipação

Antecipação
Manifestação mais precoce e aumento da gravidade de algumas doenças genéticas a cada geração

A antecipação constitui um fenômeno particular que pode influenciar a expressão de um fenótipo. Em alguns distúrbios genéticos, especialmente naqueles decorrentes de *mutações dinâmicas* (Capítulo 3), é comum observarmos a ocorrência de antecipação, fenômeno biológico que denota aumento da gravidade de uma doença genética ou sua manifestação mais precoce em sucessivas gerações. A expansão de repetições de trinucleotídios no DNA tem sido fortemente associada ao fenômeno da antecipação observado na distrofia miotônica e na doença de Huntington, ambas doenças de herança autossômica dominante. Observa-se, em genealogias nas quais essas condições estão segregando, que a gravidade e o início precoce da doença estão diretamente relacionados com a instabilidade dessas trincas que tendem a aumentar de número ao serem transmitidas às novas gerações.

Pleiotropia

Pleiotropia
Efeitos fenotípicos múltiplos de um gene

É essa a denominação dada aos efeitos fenotípicos múltiplos que um gene pode apresentar, ou seja, quando a expressão de um único gene afeta diversas características do fenótipo, muitas vezes não relacionadas, dizemos que este é pleiotrópico. Um número expressivo de síndromes monogênicas é decorrente de pleiotropia gênica, como a síndrome de Marfan, na qual se observam anomalias esqueléticas, oculares e cardiovasculares em decorrência de uma mutação autossômica dominante no gene da fibrilina 1 (*FBN1*) e a síndrome de CHARGE, distúrbio autossômico dominante causado por mutações no gene *CHD7* que resulta em coloboma da íris, defeitos cardíacos, retardo do crescimento e do desenvolvimento, anomalias das orelhas, anomalias genitais e fenda labial.

Heterogeneidade genética

Heterogeneidade genética
Produção de fenótipos iguais por mecanismos genéticos diferentes

Conforme observado na seção anterior, um único gene pode ter efeitos fenotípicos múltiplos. De maneira oposta, diferentes genótipos são capazes de produzir um mesmo fenótipo e, quando isso ocorre, considera-se que esses traços são geneticamente heterogêneos. A heterogeneidade genética representa mais um fator complicador no estabelecimento das correlações genótipo-fenótipo e, na maioria das vezes, a única maneira de distinguir o fator genético que está atuando é por meio de análise molecular.

A heterogeneidade genética resulta de mutações em diferentes *loci* (*heterogeneidade de locus*), de diferentes mutações no mesmo *locus* (*heterogeneidade alélica*) ou ambas.

Heterogeneidade de *locus*
Refere-se à ocorrência de um mesmo fenótipo causado por mutações em genes diferentes

Na heterogeneidade de *locus*, determinada condição pode ocorrer em diferentes famílias em virtude de mutações em genes distintos. Entre os fenótipos causados por mutações gênicas em mais de um *locus*, temos o câncer hereditário de mama e ovário (genes *BRCA1* e *BRCA2*), a doença de Alzheimer genética (genes *PSEN1* e *PSEN2*), a retinite pigmentosa, tipos hereditários de deficiência mental, o xeroderma pigmentoso e a surdez congênita, na qual mais de 20 genes responsáveis por formas autossômicas recessivas, dominantes ou ligadas ao X estão envolvidos na perda auditiva observada em recém-natos.

Heterogeneidade alélica
Refere-se à ocorrência de um mesmo fenótipo causado por diferentes mutações em um mesmo gene

Heterozigoto composto
Indivíduo que apresenta um traço autossômico recessivo em decorrência de duas mutações diferentes em um mesmo gene

A heterogeneidade genética também pode ocorrer ao nível dos alelos de um único *locus* (heterogeneidade alélica), quando diferentes mutações em um mesmo gene resultam em um mesmo quadro clínico, constituindo importante causa de variação. A maioria dos distúrbios monogênicos é causada não por uma mutação gênica específica, mas por mutações distintas em um único gene, de maneira que dois indivíduos podem ter um mesmo fenótipo apesar de terem alelos mutados diferentes. Como exemplo, temos inúmeras mutações no *locus* da β-globina associadas à β-talassemia e centenas de mutações no gene *CFTR* que causam a fibrose cística. Algumas pessoas apresentam duas mutações diferentes no mesmo gene e são denominadas heterozigotos compostos. Acredita-se que a maioria dos indivíduos com distúrbios autossômicos recessivos sejam heterozigotos compostos, em vez de homozigotos verdadeiros, a menos que seus genitores sejam consanguíneos.

Fenocópia

Fenocópia
Traço presente em um indivíduo com genótipo normal, causado por um fator ambiental que se assemelha a uma condição genética determinada por um genótipo específico

O termo fenocópia refere-se à manifestação de um traço adquirido, não herdado, que se assemelha a um traço de origem genética, ou seja, o efeito de um gene mutante pode ser simulado pelo efeito de um agente ambiental em um indivíduo geneticamente não mutante. Um exemplo clássico de fenocópia é a *surdez congênita* que, como vimos, além de ser causada por fatores genéticos múltiplos, pode também ser decorrente de agentes ambientais, como determinados fármacos e infecções pré e neonatais, como rubéola e toxoplasmose. Nesse caso, se ficar estabelecida que a doença foi decorrente da ação de um agente ambiental, não sendo a criança portadora de qualquer mutação gênica relacionada com o fenótipo, diremos tratar-se de uma fenocópia.

Resumindo, fatores como mutação nova, mosaicismo germinativo, penetrância reduzida, expressividade variável, efeitos gênicos pleitrópicos, heterogeneidade genética, bem como a diferenciação entre um fenótipo de origem genética e outro de causa ambiental podem obscurecer a interpretação do padrão de herança em uma genealogia. Dessa maneira, não se pode deixar de considerá-los na prática da genética clínica, de maneira a não prejudicar a confiabilidade da informação fornecida aos portadores de doenças genéticas, bem como a seus familiares. Considerando o conhecimento acumulado sobre os genes humanos, em todas as situações abordadas, estudos moleculares dos genes envolvidos, quando possível, podem auxiliar o médico.

Imprinting genômico e dissomia uniparental

O genoma humano tem uma marcação epigenética diferencial em determinadas regiões cromossômicas, denominadas *imprinting genômico*, que estabelecem padrões diferentes de expressão em genes de origem materna e paterna, ou seja, o controle da transcrição de alguns genes está condicionado à origem do alelo em questão. Consequentemente, alelos em determinados *loci* se expressam de modo monossômico, ou seja, em alguns *loci* expressamos apenas o alelo materno, permanecendo o alelo paterno inativado; em outros, expressamos apenas o alelo paterno, estando o materno inativado. É crescente a relação dos distúrbios genéticos decorrentes de mutações que causam desregulação na expressão gênica por erros de *imprinting* genômico. Como exemplos, temos a síndrome de Prader-Willi e de Angelman (15q11-q13) e a síndrome de Beckwith-Wiedemann (11p15).

Imprinting genômico
Marcação epigenética diferencial que ocorre na cromatina e acarreta padrões distintos de expressão gênica

Dissomia uniparental
Refere-se ao mecanismo que faz um indivíduo herdar ambos os cromossomos de um par de homólogos de um mesmo genitor e nenhuma cópia desse cromossomo do outro genitor

Outro mecanismo que pode causar desbalanceamento da expressão gênica em *loci* imprintados é a dissomia uniparental. Normalmente, herdamos um cromossomo homólogo de cada genitor. Entretanto, em decorrência de erros de não disjunção meiótica, é possível que um indivíduo herde duas cópias do mesmo cromossomo homólogo de um único genitor e nenhuma cópia do outro genitor, ao que denominamos dissomia uniparental. Nessa situação, embora a constituição cromossômica desse indivíduo seja normal, ele poderá manifestar algum distúrbio, dependendo se o cromossomo que sofreu dissomia uniparental apresenta genes diferentemente expressos nos cromossomos materno e paterno. Em um percentual de casos, esse mecanismo origina as síndromes de Prader-Willi e de Angelman, sendo a dissomia uniparental do cromos-

somo 15 materno associada à síndrome de Prader-Willi e a dissomia uniparental paterna do 15 à síndrome de Angelman. Outros distúrbios decorrentes de erros no *imprinting* genômico por dissomia uniparental são o diabetes neonatal transiente (dissomia uniparental paterna do cromossomo 6) e a síndrome de Beckwith-Wiedemann (dissomia uniparental paterna do cromossomo 11).

Herança materna ou matrilinear

Herança mendeliana O mesmo que herança monogênica. Refere-se aos padrões de transmissão gênica que seguem as leis de Mendel

Herança materna (ou matrilinear) Padrão de herança mitocondrial no qual a transmissão da informação genética contida no mtDNA ocorre apenas pela mulher

Para que as mitocôndrias realizem corretamente suas funções na célula é necessário haver uma perfeita interação entre os produtos de genes nucleares e mitocondriais. Isso explica por que tanto mutações no nDNA quanto no mtDNA podem comprometer o funcionamento da mitocôndria, originando doenças cuja transmissão ocorre de maneira diferenciada. Se a alteração ocorrer no nDNA, será transmitida por herança mendeliana. Se ocorrer no mtDNA, apresentará um tipo peculiar de transmissão, a herança materna (ou matrilinear), significando que esses traços são transmitidos exclusivamente pelas mulheres, embora acometam ambos os sexos.

Considerando que a transmissão de fenótipos ligados a mutações no genoma mitocondrial ocorre exclusivamente pela linhagem materna, um homem que herde de sua mãe uma mutação mitocondrial não a transmitirá à sua prole. Desse modo, traços mitocondriais monogênicos não segregam de forma mendeliana clássica. Na Figura 5.10, apresentamos um heredograma mostrando o padrão de herança matrilinear. Observe que tanto homens como mulheres podem apresentar o traço, embora somente as mulheres possam transmiti-lo.

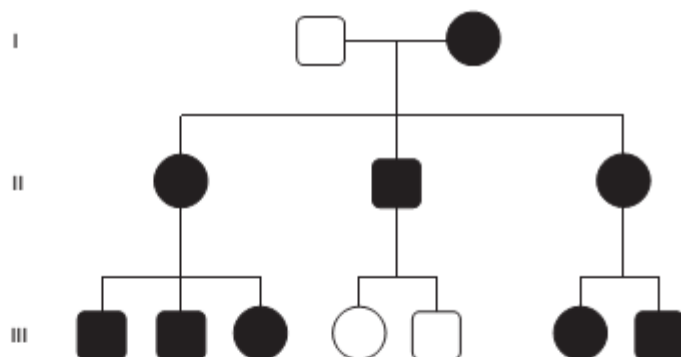



Figura 5.10 Heredograma que exemplifica a herança materna ou matrilinear.

Como exemplos de doenças causadas por alterações em genes localizados no DNA mitocondrial, temos: neuropatia óptica hereditária de Leber (LHON), que leva a perda súbita e irreversível da visão, síndrome de Kearns-Sayre (oftalmoplegia, degeneração pigmentar da retina e cardiomiopatia), epilepsia mioclônica com fibras vermelhas anfractuosas (MERRF), encefalomiopatia, acidose láctica e episódios semelhantes a acidente vascular cerebral (MELAS), síndrome de neuropatia, ataxia e retinite pigmentosa (NARP) e síndrome de Pearson (insuficiência pancreática e da medula óssea).

RESUMO

- A transmissão dos caracteres monogênicos, também denominados mendelianos, segue o modelo de segregação descrito por Mendel
- OMIM é um banco de dados *on-line* que reúne extensa informação sobre genes e fenótipos monogênicos em humanos
- Heredograma ou árvore genealógica é a representação gráfica de uma história familiar pela qual é possível acompanhar a transmissão de um traço pelas gerações
- O grau de consanguinidade se expressa pelo coeficiente de endocruzamento (F), definido como a probabilidade de um

indivíduo receber, em determinado *locus*, dois alelos idênticos por herança

- A ausência quase total de homologia entre os cromossomos sexuais humanos X e Y nos possibilita afirmar que os genes localizados no X apresentam padrões de herança diferentes nos dois sexos, em virtude de a mulher apresentar 2 cromossomos X e o homem apenas 1. No sexo feminino, traços ligados ao X mantêm as relações de dominância e recessividade, ao passo que, no sexo masculino, eles se expressam em hemizigose
- Alguns fenótipos autossômicos apresentam manifestação restrita ao sexo masculino ou feminino e existem traços que, embora se expressem em ambos os sexos, sofrem a influência do sexo
- São características da herança autossômica dominante: transmissão vertical do fenótipo, sem pular gerações; o indivíduo afetado tem um dos genitores afetados; ambos os sexos podem apresentar e transmitir o traço; afetados transmitem a condição para metade de sua prole
- São características da herança autossômica recessiva: transmissão horizontal do fenótipo; pais de indivíduos afetados, em geral, são heterozigotos e transmitirão o traço para 1/4 de sua prole; homens e mulheres têm igual probabilidade de apresentar e transmitir o traço; elevada taxa de consanguinidade entre os genitores
- São características da herança recessiva ligada ao X: em geral, apenas o sexo masculino é afetado; o traço não é transmitido de homem para homem; a transmissão, em geral, se dá por mulheres heterozigotas normais; os homens afetados transmitem o gene mutante para todas as suas filhas, que serão portadoras
- São características da herança dominante ligada ao X: os homens afetados transmitem o traço para todas as suas filhas e para nenhum de seus filhos; as mulheres afetadas heterozigotas transmitem a condição para metade de seus filhos de ambos os sexos; homens e mulheres podem apresentar o traço, sendo este mais prevalente nas mulheres
- São características da herança ligada ao Y (holândrica): somente homens são afetados e ocorre transmissão exclusiva de homem para homem
- Em alguns distúrbios dominantes, uma proporção significativa de indivíduos acometidos resulta de mutações novas
- O mosaicismo germinativo ou gonadal contribui para o aparecimento de novos traços genéticos na prole de indivíduos normais
- Define-se penetrância como a probabilidade do efeito de um gene se manifestar ou não ao nível de fenótipo. Traços que são expressos em 100% dos indivíduos portadores da mutação têm penetrância completa e aqueles cuja expressão fenotípica ocorre apenas em parte dos indivíduos apresentam penetrância reduzida ou incompleta
- Doenças genéticas de início tardio apresentam penetrância relacionada com a idade
- É comum, em traços autossômicos dominantes, observarmos variação nos efeitos fenotípicos produzidos por um mesmo gene em diferentes indivíduos, ao que denominamos expressividade variável
- Em alguns distúrbios genéticos, verificamos a ocorrência de antecipação, definida como o aumento da gravidade de uma doença ou sua manifestação mais precoce em sucessivas gerações
- Um número expressivo de síndromes monogênicas é decorrente de pleiotropia gênica, que se refere aos efeitos fenotípicos múltiplos que um gene pode apresentar
- Diferentes genótipos são capazes de produzir um mesmo fenótipo e, quando isso ocorre, considera-se que esses traços são geneticamente heterogêneos. A heterogeneidade genética pode resultar de mutações em diferentes *loci* (heterogeneidade de *locus*), de diferentes mutações no mesmo *locus* (heterogeneidade alélica) ou ambas
- O termo fenocópia refere-se à simulação do efeito de um gene mutante por um agente ambiental em um indivíduo geneticamente não mutante
- Fatores como mutação nova, mosaicismo germinativo, penetrância reduzida, expressividade variável, efeitos gênicos pleitrópicos, heterogeneidade genética e fenocópia podem dificultar a interpretação do padrão de herança em uma genealogia
- Alelos em determinados *loci* se expressam de modo monossômico (somente o de origem materna, permanecendo o alelo de origem paterna inativado ou vice-versa)
- Embora, normalmente, o ser humano herde um cromossomo homólogo de cada genitor, é possível, em decorrência de erros de não disjunção meiótica, que um indivíduo herde duas cópias do mesmo cromossomo homólogo de um único genitor e nenhuma cópia do outro genitor, ao que denominamos dissomia uniparental
- A transmissão de fenótipos ligados a mutações no genoma mitocondrial ocorre exclusivamente pela mulher (herança materna ou matrilinear), de modo que um homem, ao herdar de sua mãe uma mutação mitocondrial, não a transmitirá a seus descendentes

AUTOAVALIAÇÃO

5.1 Conceitue traços dominantes, codominantes e recessivos.

5.2 Você acha que, com base nas orientações recebidas, seria capaz de montar um heredograma exemplificando a transmissão de um traço em sua família?

5.3 Qual o significado do termo probando?

5.4 Quais critérios você identifica como sendo característicos da herança autossômica dominante? Compare-os com aqueles que definem a herança autossômica recessiva.

5.5 A partir do que foi apresentado, você seria capaz de identificar as principais diferenças entre os padrões de transmissão monogênica ligados ao X e autossômicos?

5.6 Qual é a relação entre casamento consanguíneo e traços autossômicos recessivos?

5.7 Defina coeficiente de endocruzamento.

5.8 Explique a razão pela qual a herança de traços ligados ao X ocorre de maneira diferenciada nos dois sexos.

5.9 Explique por que fenótipos recessivos ligados ao X geralmente só ocorrem em homens e traços dominantes ligados ao X são mais frequentemente encontrados nas mulheres.

5.10 Conceitue penetrância, expressividade variável, antecipação, pleiotropia e fenocópia.

5.11 Comente como o mosaicismo gonadal pode levar ao surgimento de um distúrbio genético.
5.12 Diferencie heterogeneidade de *locus* de heterogeneidade alélica.
5.13 Qual é a característica marcante da herança holândrica?
5.14 Diferencie herança ligada ao sexo de herança limitada ou influenciada pelo sexo.
5.15 Comente a afirmativa: “Nem toda doença genética é congênita.”
5.16 Elabore um resumo sobre a relação do *imprinting* genômico e da dissomia uniparental com o desbalanceamento dos padrões de expressão gênica em humanos.
5.17 Conceitue herança materna ou matrilinear.

BIBLIOGRAFIA

BENNET, R.L.; STEINHAUS, K.H.; UHRICH, S.B. *et al.* Recommendation for standardized human pedigree nomenclature. *Am J Hum Genet* 1995; 56:745-752.

Online Mendelian Inheritance in Man (OMIM). Disponível em http://www.ncbi.nlm.nih.gov/omim/. Acesso em 01/02/2012.

NASSEH, I.E.; TENGAN, C.H.; KIYOMOTO, B.H.; GABBAI, A. A. Doenças mitocondriais. *Rev Neurociências* 2001; 9(2): 60-69.

NUSSBAUM, R.L.; MCINNES, R.R.; WILLARD, H.F. *Thompson & Thompson genética médica.* 7ª ed., Rio de Janeiro: Elsevier, 2007.

PASSARGE, E. *Genética. Texto e atlas.* 2ª ed., Porto Alegre: Artmed, 2004.

ROSS, M.T.; GRAFHAM, D.V.; COFFEY, A.J. *et al.* The DNA sequence of the human X chromosome. *Nature* 2005; 434: 325-337.

SKALETSKY, H; KURODA-KAWAGUCHI, T.; MINX, P.J. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes *Nature* 2003; 423: 825-837.

TURNPENNY, P.D.; ELLARD, S. *Emery genética médica.* 13ª ed., Rio de Janeiro: Elsevier, 2009.

WALTER, J.; PAULSEN, M. Imprinting and disease. *Semin Cell Dev Biol* 2003; 14(1): 101-110.

6

Herança Poligênica e Multifatorial

Objetivos de estudo

Ser capaz de conceituar e distinguir os tipos de herança: monogênica, oligogênica, poligênica e multifatorial

Conhecer os tipos de variação fenotípica presentes nas populações, bem como a base genética dessa variabilidade

Compreender as interações possíveis entre os poligenes na variação quantitativa

Saber diferenciar herança poligênica contínua de herança poligênica dicotômica

Reconhecer de que maneira os fatores ambientais podem contribuir para a variação fenotípica nas populações

Saber o que representam a plasticidade fenotípica e a norma de reação

Ser capaz de distinguir fenótipos quantitativos contínuos, categóricos e dicotômicos

Conhecer os parâmetros estatísticos relativos à distribuição normal das variáveis contínuas (curva gaussiana)

Perceber a diferença entre traços poligênicos dicotômicos e fenótipos quantitativos monogênicos

Explicar a variação fenotípica quantitativa dicotômica com base no modelo limiar ou de propensão

Perceber como a variabilidade fenotípica é composta

Compreender o conceito de herdabilidade e a distinção entre herdabilidade em sentido amplo e em sentido restrito

Conhecer o método de gêmeos e sua aplicabilidade

Ter noção da aplicação dos estudos de associação genômica na identificação dos QTL

Perceber o significado dos SNP como importantes marcadores da variação genética individual e sua relação com as doenças complexas humanas

Conceitos-chave do capítulo

Desvio padrão (s)
Dominância
Característica multifatorial
Característica poligênica categórica
Característica poligênica contínua
Chip de DNA
Codominância
Epistasia
Fenótipo
Gêmeos dizigotos (fraternos ou bivitelinos)
Gêmeos monozigóticos (idênticos ou univitelinos)
Genoma
Genótipo
Herança digênica
Herança monogênica
Herança multifatorial
Herança oligogênica
Herança poligênica
Herança quantitativa
Herdabilidade
Herdabilidade em sentido restrito (h^2)
Herdabilidade em sentido amplo (H^2)
Loci de característica quantitativa (QTL)
Média
Método de gêmeos
Modelo limiar ou de propensão
Norma de reação
Plasticidade fenotípica
Poligenes
Poligenes deletérios
Polimorfismo de nucleotídio único (SNP)
Variação dicotômica
Variação fenotípica contínua
Variação fenotípica descontínua
Variância ambiental
Variância fenotípica total
Variância genética aditiva
Variância genotípica

Introdução

Genótipo
Constituição genética de uma célula ou de um organismo

Fenótipo
Característica observada em um organismo, determinada pela interação de seu genótipo com o ambiente em que se expressa

Nem sempre a relação genótipo-fenótipo ocorre de modo simples e previsível como nas características monogênicas abordadas no Capítulo 5. Como veremos a seguir, inúmeros traços herdados são mais complexos em relação aos fatores genéticos a eles associados. Na realidade, a maioria da *variabilidade fenotípica* observada em todos os organismos, animais ou vegetais não resulta da expressão de um gene único que produz um efeito fenotípico nítido, mas decorre da *ação combinada de alguns* ou de *vários genes em diferentes loci* espalhados no genoma, ou seja, é de natureza *poligênica*.

A segunda metade do século XIX foi um período marcado por intensa investigação científica, a qual buscava explicar a variabilidade genética existente nas populações naturais. Nessa épo-

Herança poligênica
Refere-se à herança de características fenotípicas cuja expressão depende da ação conjunta de vários genes, cada um produzindo um pequeno efeito sobre o traço

Herança multifatorial
Refere-se à herança de fenótipos determinados por interações entre vários fatores genéticos e ambientais

ca, a complexidade da variação fenotípica dos *traços poligênicos* já era percebida por Francis Galton (1889) em suas observações sobre a distribuição desses caracteres entre genitores e seus descendentes, mesmo antes de ele ter conhecimento das leis da hereditariedade propostas por Gregor Mendel. Entretanto, a compreensão sobre os fundamentos dessa variabilidade fenotípica só ocorreu no século seguinte, após os estudos do estatístico Ronald Aymler Fisher (1918), cujo trabalho em muito contribuiu para a genética quantitativa e de populações.

Nas características de herança poligênica, além dos efeitos e interações entre os múltiplos genes envolvidos, *fatores ambientais* também podem contribuir para a variação observada nessas características, constituindo o que denominamos herança multifatorial. Como exemplos de fenótipos resultantes de interações complexas entre genes e ambiente, destacamos estatura, peso, cor da pele e do cabelo, inteligência, atividade metabólica e inúmeras doenças complexas.

Estima-se que mais de 70% das doenças humanas herdadas, inclusive as de manifestação tardia, sejam multifatoriais, com prevalência dessas condições 30 vezes maior que a incidência dos distúrbios monogênicos.

■ Variação fenotípica nas populações | Características qualitativas e quantitativas

Variação monogênica discreta (qualitativa)

A variabilidade populacional relativa a determinado fenótipo pode ser causada pela expressão de diferentes alelos de um único gene, constituindo o que denominamos *variação monogênica ou mendeliana discreta*. Para relembrar, em seus experimentos com ervilhas, Mendel estudou características simples que envolviam um pequeno número de fenótipos nitidamente distintos (como a forma das sementes, lisas ou rugosas), resultantes da segregação de alelos de apenas um gene. Assim como os traços hereditários analisados por Mendel, outros fenótipos, também determinados pela expressão de genes únicos, contribuem para a variação fenotípica descontínua observada nas populações. Eles constituem um grupo de características que exibem *variabilidade fenotípica discreta*, de natureza *qualitativa*.

Variação fenotípica descontínua
Padrão de variabilidade relativa a uma característica cujos fenótipos se apresentam em categorias nitidamente distintas e independentes

A nítida contribuição que o produto de um gene único exerce no fenótipo possibilita que, em relação à característica específica, os indivíduos de uma população possam ser classificados em um pequeno número de categorias independentes, que não mostrem ligação entre si. Traços qualitativos são bem exemplificados pelos diferentes grupos sanguíneos eritrocitários humanos. Nos sistemas ABO, MN e Rh, a capacidade de aglutinação das hemácias na presença dos antissoros específicos possibilita que as pessoas sejam distintamente classificadas de acordo com os antígenos que apresentam. Observe a Tabela 6.1, relativa ao *locus* do grupo sanguíneo MN, localizado no cromossomo 4 humano. A combinação de dois alelos codominantes, L^M e L^N, determinantes da presença dos antígenos M e N nas hemácias, resulta em três fenótipos distintos (M, MN e N), de maneira que a união de um casal, ambos heterozigotos (MN), produzirá prole nas seguintes proporções: 1 L^ML^M:2 L^ML^N:1 L^NL^N.

Como se pode constatar nesse exemplo de *variação fenotípica discreta* ou *descontínua*, cada indivíduo, em decorrência de seu genótipo, pertence a um subgrupo de pessoas com fenótipos mutuamente exclusivos, o que torna possível estabelecer uma classificação qualitativa. Por não sofrerem efeitos ambientais significativos, a transmissão de características monogênicas qualitativas pode ser diretamente acompanhada e predita com base na análise de segregação do traço por meio das gerações em uma genealogia, como vimos no Capítulo 5.

■ **Tabela 6.1** Fenótipo qualitativo humano que exibe variação descontínua, exemplificado pelo grupo sanguíneo MN, condicionado por herança monogênica.

Genótipo	Fenótipo (grupo sanguíneo)
L^ML^M	M
L^ML^N	MN
L^NL^N	N

A variação fenotípica de origem monogênica está diretamente relacionada com a nítida contribuição que o produto de um único gene exerce no fenótipo.

Poligenes
Genes envolvidos na herança quantitativa ou poligênica

Variação fenotípica contínua
Tipo de variação observada em características quantitativas, na qual se verifica um contínuo de fenótipos entre os indivíduos, não sendo observadas classes fenotípicas nitidamente distintas, como ocorre na variação descontínua

Herança quantitativa
O mesmo que herança poligênica

Herança oligogênica
Refere-se à herança de fenótipos que resultam da ação combinada de poucos genes (em geral, 2 a 15)

Herança digênica
Tipo de herança oligogênica, na qual o fenótipo resulta de combinações de alelos em 2 genes

Variação contínua (quantitativa)

Embora o exemplo motrado na Tabela 6.1 seja representativo de uma *característica qualitativa de herança mendeliana simples*, é importante salientar que a maioria dos fenótipos observados nas populações naturais não segrega de maneira monogênica, o que significa que a variabilidade fenotípica observada em muitos traços hereditários não é consequência apenas da expressão diferencial dos alelos de um único gene, mas resulta da *ação combinada de genes em diferentes loci* (poligenes), podendo também sofrer a influência do ambiente. Os produtos dos diferentes genes interagem, cada um com pequeno efeito cumulativo na determinação da característica, o que causa variação fenotípica contínua. Desse modo, os *fenótipos poligênicos* são de herança quantitativa e, diferentemente dos monogênicos, exibem variação contínua nas populações, isto é, se distribuem de maneira gradual entre dois extremos.

Presume-se que a maioria dos traços quantitativos resulte da ação conjunta de muitos genes (em geral, mais de 15), cada um exercendo um pequeno efeito fenotípico, como é o caso do diabetes tipo 1 de manifestação juvenil, decorrente da complexa interação de um *locus* principal (HLA) com cerca de outros 20 *loci*. No entanto, alguns fenótipos conhecidos são resultantes da atividade combinada de poucos genes (2 a 15), constituindo o que denominamos herança oligogênica. São exemplos de características oligogênicas alguns tipos de surdez não sindrômica e de retinite pigmentosa e a síndrome de Bardet-Biedl, esta caracterizada por distrofia da retina, obesidade, polidactilia, defeitos renais, déficit cognitivo e testículos pouco desenvolvidos. Um tipo específico de retinite pigmentosa, doença que leva à degeneração da retina com consequente perda da visão, resulta de mutações conjuntas em dois genes (*RDS* e *ROM1*) e, portanto, constitui um exemplo de herança digênica. Esse também é o caso de um tipo de surdez não sindrômica, causada por mutações nos genes *GJB2* e *GJB6*.

Embora os princípios mendelianos de segregação e distribuição gênica também se apliquem a traços quantitativos, nota-se que o padrão de *herança poligênica* é bem mais complexo, pois a segregação dos genes não é evidente.

Herança monogênica
O mesmo que herança mendeliana. Refere-se aos padrões de transmissão de alelos de um único gene, presente no genoma nuclear de um organismo diploide

Tomemos como exemplo dois fenótipos quantitativos, facilmente observados nos indivíduos de uma população, como a estatura e a cor dos cabelos. Como prever qual será a altura ou a cor dos cabelos dos filhos de um casal cujo homem tenha 1,84 m e a mulher, 1,62 m de estatura, sendo ela naturalmente loira e ele de cabelos castanhos? Na verdade, mesmo sem conhecimento mais profundo de genética, é fácil concluir que para essas características, embora possamos inferir a existência de similaridades entre os genitores e sua prole, pois quanto maior o grau de parentesco maior a analogia fenotípica, não temos como precisar quais os fenótipos possíveis nos descendentes desse casal. Ou seja, mesmo conhecendo o fenótipo dos genitores, não temos como estabelecer classes fenotípicas nitidamente distintas para sua prole, como ocorreu em relação ao grupo sanguíneo MN. Se esse casal tiver 10 filhos, provavelmente todos terão alturas e tonalidades de cabelo diferentes. Nessa situação, a *análise do heredograma*, amplamente utilizada na herança monogênica, não contribui para as previsões sobre a transmissão desses fenótipos quantitativos, sendo necessário dispor de dados populacionais com os quais inferências estatísticas sejam derivadas.

Por serem os traços quantitativos condicionados pela ação de múltiplos genes, cada um segregando independentemente, com um efeito específico, o fenótipo resultante será produto da ação dos vários alelos nos diferentes *loci*. À medida que o número de *loci* gênicos envolvidos aumentar, também aumentará o número de fenótipos resultantes das interações alélicas. Isso pode ser facilmente percebido na Figura 6.1. Ela apresenta um exemplo hipotético de variação fenotípica referente a uma característica condicionada pela interação de 2, 3, 4 ou 5 genes, representados pelos alelos *F/f*, *G/g*, *H/h*, *I/i* e *J/j*. É importante salientar que se tratam de genes não ligados, ou seja, que segregam de maneira independente um do outro. Observe a distribuição das classes fenotípicas, resultantes das combinações gênicas possíveis, esperadas na descendência de dois indivíduos heterozigotos:

- Se o traço resultar da ação de pares de alelos em 2 *loci*, teremos a presença de 5 fenótipos distribuídos na frequência de 1:4:6:4:1
- Se o traço for condicionado por pares de alelos em 3 *loci*, teremos 7 classes fenotípicas na frequência de 1:6:15:20:15:6:1
- Se o traço for determinado pela combinação de pares de alelos em 4 genes, teremos 9 classes fenotípicas (1:8:28:56:70:56:28:8:1)
- No caso de o traço ser resultante da ação de pares de alelos em 5 *loci*, teremos 11 classes fenotípicas (1:10:45:120:210:252:210:120:45:10:1).

Torna-se evidente, portanto, que o aumento do número de fatores genéticos (poligenes) que contribuem para um traço quantitativo resulta em um número crescente de classes fenotípicas e em uma distribuição cada vez mais uniforme da variabilidade fenotípica, facilmente observada nas populações. Note que essa uniformidade da distribuição já pode ser percebida

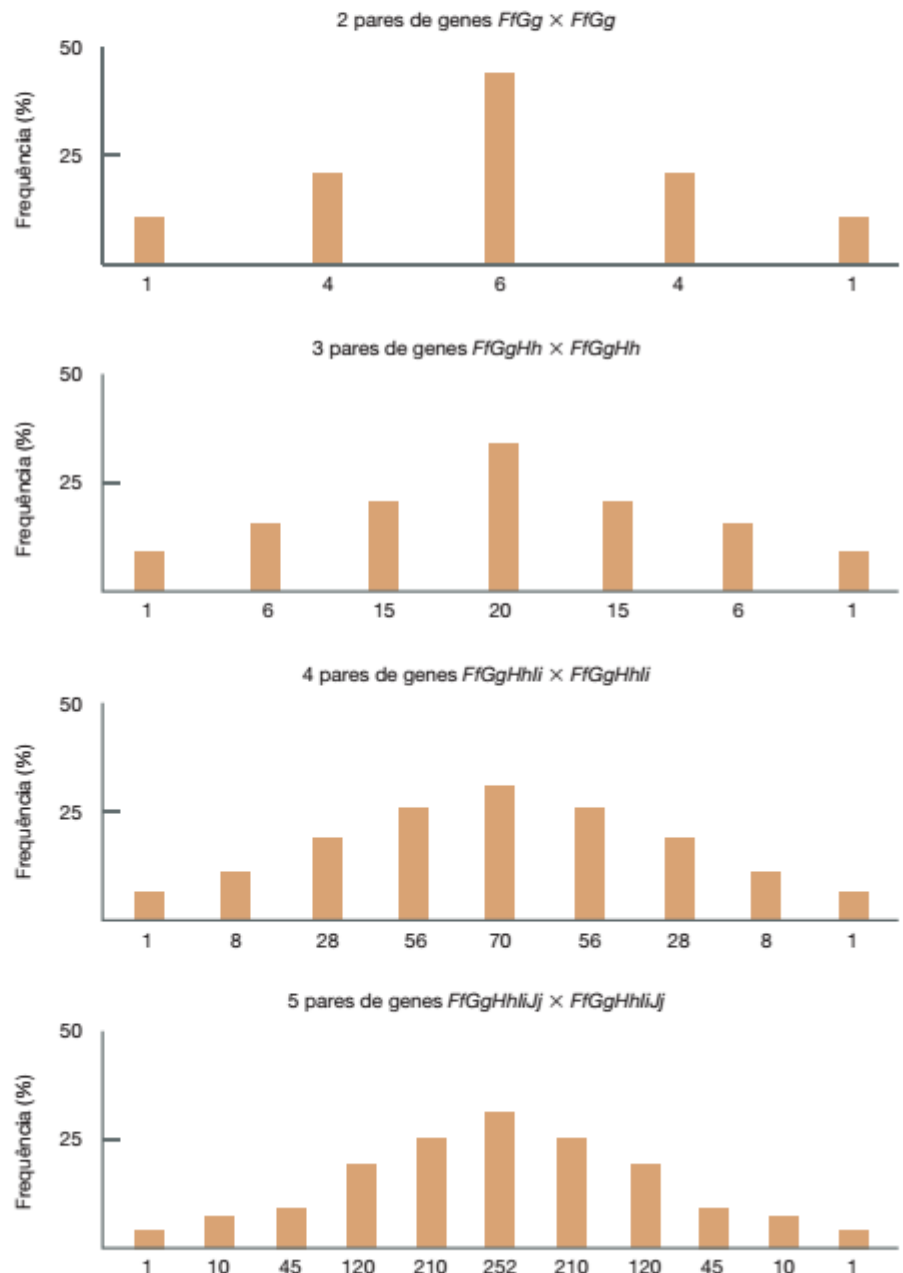



Figura 6.1 Distribuição esperada de classes fenotípicas, representadas pelas barras do histograma, na descendência de indivíduos heterozigotos para um traço poligênico, condicionado por 2, 3, 4 ou 5 genes não ligados.

mesmo quando apenas 3 *loci* contribuem para o fenótipo. Os *loci* que contribuem para um traço quantitativo, denominados loci de característica quantitativa (QTL, ***quantitative trait loci***), serão abordados de modo mais abrangente mais adiante.

Loci de característica quantitativa (QTL, *quantitative trait loci*)
Genes que contribuem para um traço quantitativo

Ainda em relação ao componente genético, devemos destacar que a variabilidade fenotípica dos traços quantitativos é resultante não apenas da combinação entre os diferentes genes, mas também de complexas ***interações alélicas intralocus*** (dominância ou codominância) e ***entre loci*** (epistasia ou ***efeitos aditivos***), em grande parte desconhecida, que contribuem para a diversidade de fenótipos observados. É importante notar que as interações intergênicas nem sempre são de natureza aditiva, de modo a contribuição dos genes para determinada característica pode ocorrer de maneira desigual. Alguns genes podem ter efeito mais significativo na determinação do traço do que outros.

Dominância
Característica expressa fenotipicamente em heterozigotos para o alelo determinante do traço

Codominância
Característica em que os efeitos de dois alelos distintos presentes em um *locus* no par de cromossomos homólogos são observados no fenótipo

Para tornar mais claro o entendimento das características quantitativas, podemos agrupá-las em duas categorias: ***poligênicas contínuas*** e ***poligênicas dicotômicas***. Antes de abordá-las, no entanto, faremos algumas considerações sobre a contribuição do ambiente na variabilidade dos traços quantitativos.

Epistasia
Refere-se a interações não recíprocas entre produtos de genes não alélicos, de modo que um gene em um *locus* pode interferir na expressão de outro gene situado em outro *locus*

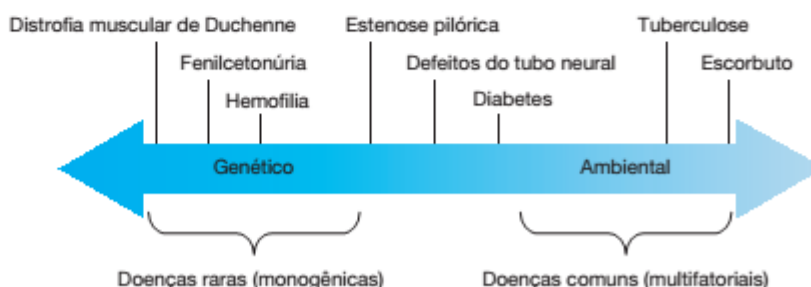
A influência do ambiente na variação fenotípica poligênica | Herança multifatorial

Além dos efeitos que diferentes genes somam para a variação de um fenótipo poligênico, ***fatores ambientais*** diversos também podem contribuir para essa variabilidade, sendo o traço em questão denominado ***traço multifatorial***. Na verdade, a variação fenotípica quantitativa observada nas populações naturais é, em sua maioria, de natureza multifatorial, ou seja, determinada não apenas pelo genótipo, mas pela ***interação entre fatores genéticos múltiplos com o ambiente*** ao qual o organismo está exposto nas diferentes fases de seu desenvolvimento, incluindo hábitos alimentares e estilo de vida (tabagismo, consumo de bebidas alcoólicas, atividade física, exposição a agentes tóxicos). A altura é um bom exemplo de característica multifatorial, por ser um traço genético condicionado por vários genes, também influenciado pela dieta e atividade física.

Característica multifatorial
Aquela que resulta da combinação de vários fatores genéticos e ambientais

A complexidade dos fenótipos multifatoriais relaciona-se ao fato de eles serem determinados por poligenes e sofrerem a influência do ambiente. Dessa maneira, para esses fenótipos, membros de uma mesma família podem ser semelhantes tanto por compartilharem alelos em comum quanto por estarem expostos aos mesmos agentes ambientais.

Cabem aqui algumas considerações mais gerais sobre a ***relação genótipo-ambiente-fenótipo***. Por definição, todo fenótipo é resultado da interação entre fatores genéticos e ambientais, embora a contribuição do genótipo e do ambiente ocorra de modo diferenciado nos diferentes traços. Isso pode ser facilmente percebido nas doenças humanas, que se distribuem entre dois opostos, como mostrado na Figura 6.2. Observe que em algumas doenças, como a distrofia muscular de Duchenne (herança monogênica recessiva ligada ao X), o fator ambiental não contribui de maneira significativa para a manifestação fenotípica do traço, e, nessa categoria, incluem-se outras doenças monogênicas, embora sejam raras aquelas que sofrem pouca ou nenhuma influência do ambiente. No entanto, para outras enfermidades, a exposição a fatores ambientais específicos é altamente determinante do aparecimento da doença, como no caso das doenças infecciosas. A maioria dos distúrbios, no entanto, distribui-se entre esses extremos. Verificamos, portanto, que quanto maior a influência do componente genético em relação aos

Figura 6.2 Contribuição de fatores genéticos e ambientais nas doenças humanas. Observe a distribuição entre dois opostos: a manifestação de algumas doenças é determinada por mutações em genes únicos e sofrem pouca ou nenhuma influência do ambiente. A manifestação de outras doenças está diretamente relacionada à exposição a fatores ambientais específicos. A maioria dos distúrbios se distribui entre esses dois extremos.

fatores ambientais na variabilidade fenotípica de um traço, mais genético esse traço será e, de maneira oposta, quanto menor for a influência do genótipo sobre a variabilidade fenotípica, menos genético ele será.

Plasticidade fenotípica
Propriedade inerente aos organismos vivos que lhes possibilita responder de modo diferenciado a mudanças no meio ambiente

Norma de reação
Espectro de fenótipos que podem resultar de um genótipo específico em ambientes diversos

Genoma
Informação genética total presente na célula de um organismo

Desse modo, dizer que um traço é ou não genético é apenas uma maneira relativa de expressar o quanto determinado genótipo contribui para o fenótipo.

Quando pensamos na influência de *fatores ambientais* sobre a *variação fenotípica poligênica*, é comum imaginarmos o ambiente como um agente aleatório que desvirtua a expressão do genótipo. Na verdade, todos os organismos apresentam uma plasticidade fenotípica, o que significa que eles são capazes de responder diferentemente a mudanças no meio, de modo que determinado genótipo pode, em ambientes diversos, originar um espectro de fenótipos, referido como norma de reação.

Todo genoma é sensível às variações ambientais e inúmeros estudos buscam compreender de que maneira esses fatores, mais precisamente hábitos de vida, interferem no genótipo, modificando a atividade gênica e ocasionando determinados fenótipos. Em algumas *doenças multifatoriais*, como a hipertensão, é fácil identificar alguns agentes ambientais de risco (obesidade, altos níveis de sódio na dieta, baixa atividade física, entre outros). Entretanto, na maioria dos traços multifatoriais, muito pouco ainda se sabe sobre os fatores ambientais que representam risco e sobre aqueles que são protetores. Essa compreensão é de real importância, pois abrirá novos caminhos para o delineamento de intervenções preventivas para inúmeros distúrbios.

Características poligênicas contínuas

Característica poligênica contínua
Característica quantitativa que exibe variação fenotípica contínua, ou seja, os fenótipos se distribuem gradualmente por um espectro sucessivo de medidas

São aquelas que se distribuem de maneira contínua na população, ou seja, nas quais se observa a presença de fenótipos sucessivos, sem uma separação nítida de categorias. As características poligênicas contínuas são representadas por variáveis que podem ser *quantificadas ou medidas*. Traços quantitativos humanos, que exibem ampla variação contínua, são bem exemplificados pelo peso corporal, altura, pressão arterial, inteligência, circunferência cefálica, o que significa que eles se distribuem gradualmente por um espectro sucessivo de valores. Se, aleatoriamente, medirmos a estatura em uma amostra de indivíduos representativa de uma população e lançarmos as medidas em um gráfico, obteremos uma infinidade de classes fenotípicas, sendo o número de fenótipos diretamente proporcional à precisão dos instrumentos de medição utilizados.

Característica poligênica categórica
Fenótipo poligênico que apresenta um número definido de classes fenotípicas e não uma infinidade, como observado nas características quantitativas contínuas

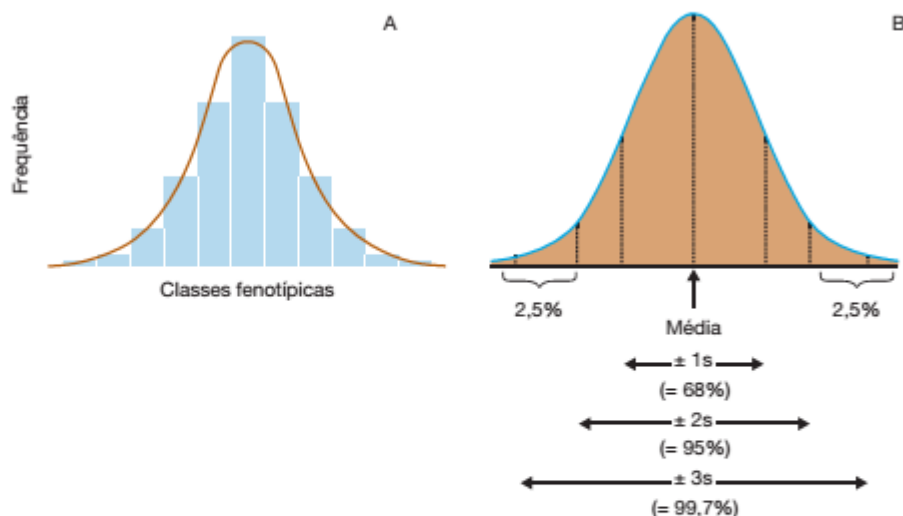
De modo alternativo, alguns fenótipos poligênicos, denominados *categóricos*, podem ser facilmente separados em classes. Nos *fenótipos categóricos*, a variação fenotípica está relacionada com valores absolutos, determinados por contagem, como o tamanho da prole, o número de sementes por vagem ou de grãos nas espigas de milho, ou mesmo o número de cerdas presentes no abdome da *Drosophila*, entre outros exemplos. Em geral, uma *Drosophila* pode apresentar 14 a 24 cerdas abdominais e uma vagem pode conter 2 a 6 sementes, mas nenhuma *Drosophila* terá 15,7 cerdas e nenhuma vagem apresentará 4,8 sementes. Desse modo, as características poligênicas categóricas, embora sejam quantitativas, não apresentam uma infinidade de fenótipos.

Média
Média aritmética

À medida que o número de genes relacionados com um traço quantitativo contínuo aumenta, verifica-se que a distribuição das classes fenotípicas cada vez mais se assemelha a uma *curva normal* ou *gaussiana*, reconhecida por sua forma simétrica, semelhante a um sino (Figura 6.3). A simetria observada na curva normal ocorre em função dos valores médios serem os mais frequentes e os valores extremos mais raros, ou seja, a maioria dos *fenótipos quantitativos* se distribui ao redor da média.

Desvio padrão (s)
Medida estatística, representada pela raiz quadrada da variância, que expressa a variação de um conjunto de valores ao redor da média

Observe na Figura 6.3B que o desvio padrão *da distribuição normal*, definido como a *raiz quadrada da variância* e representado pela letra "s", determina a dispersão da variável ao redor da média ($\bar{x}$). *Em uma distribuição normal, aproximadamente, 68%, 95% e 99,7% dos valores de uma variável encontram-se no intervalo delimitado pela média mais ou menos 1, 2 ou 3 desvios padrão, respectivamente.* Desse modo, 95% das medidas estarão distribuídas no intervalo entre dois desvios padrão acima e abaixo da média, o que denota que, ao se empregar

Figura 6.3 Distribuição normal ou gaussiana. (A) Histograma de distribuição de uma característica quantitativa contínua, em determinada amostra, que destaca a sobreposição da curva normal ajustada às medidas. (B) Dispersão da variável ao redor da média: proporções aproximadas de 68% ($\bar{x}$ ± 1s); 95% ($\bar{x}$ ± 2s) e 99,7% ($\bar{x}$ ± 3s).

esse nível de significância, a probabilidade de encontrar uma medida fora desse intervalo é de 5% (2,5% em cada extremo da distribuição).

Como regra, na prática médica, valores extremos de um traço quantitativo, situados nas extremidades da curva normal, merecem especial atenção. Já nos primeiros anos de vida, as crianças são acompanhadas para altura, peso e perímetro cefálico, usando-se, para fins comparativos, curvas de distribuição padrão dessas variáveis em uma população específica, organizadas por sexo e faixa etária. É comum na prole de indivíduos que apresentam fenótipos extremos observarmos uma tendência à ***regressão à média***, de maneira que, tomando-se o exemplo da altura em filhos de casais muito baixos ou muito altos, verificamos que eles tendem a ter estatura próxima à média da população, não sendo, necessariamente, tão altos ou baixos como seus pais.

Os fenótipos poligênicos, diferentemente dos monogênicos, apresentam uma variação contínua na população.

Características poligênicas dicotômicas | Modelo limiar ou de propensão

Variação dicotômica
O mesmo que variação fenotípica descontínua ou binária

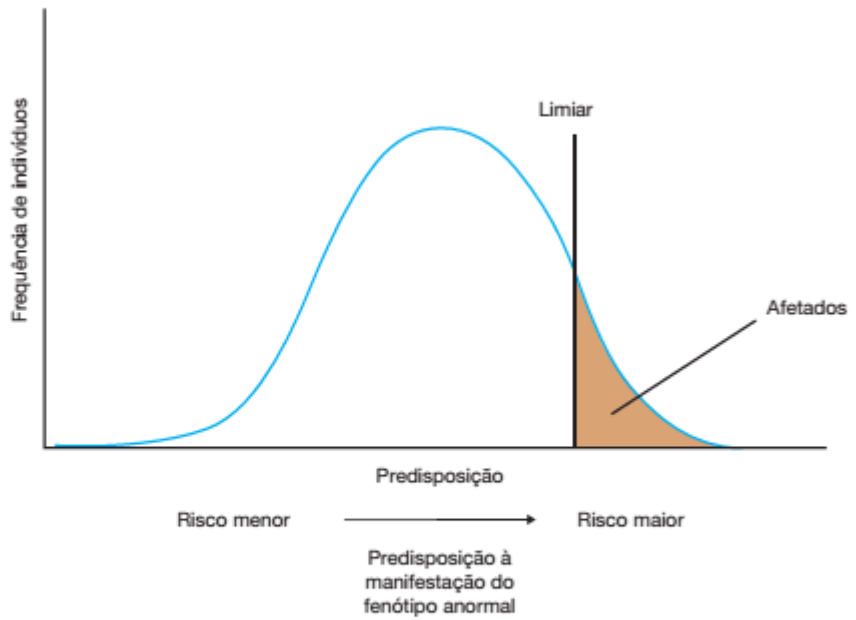
Modelo limiar ou de propensão
Tal modelo explica a variação fenotípica quantitativa dicotômica, com base na proposição de existência de um limiar que separa os indivíduos em duas categorias: os que manifestam o fenótipo anormal e os que não manifestam

Poligenes deletérios
Genes que não desempenham adequadamente suas funções na célula e causam distúrbios ao organismo

Como vimos, a variação fenotípica contínua é um atributo dos caracteres poligênicos. No entanto, alguns traços, embora de origem multifatorial e normalmente distribuídos, podem se apresentar de maneira ***dicotômica*** na população. É possível compreender a variação dicotômica desses fenótipos com base no modelo limiar ou de propensão, que estabelece a existência de um ***limiar crítico na curva de distribuição normal*** que distingue duas categorias fenotípicas nitidamente opostas. Ao observar a curva de propensão de um fenótipo quantitativo dicotômico hipotético na população (Figura 6.4), note que o limiar evidencia dois grupos de indivíduos: aqueles não afetados, nos quais o traço está ausente, e aqueles que expressam o fenótipo anormal (área sombreada da curva de distribuição fenotípica), embora exista nítida variabilidade fenotípica contínua em ambos os grupos (afetados e não afetados).

Podemos inferir que indivíduos afetados, com fenótipos acima do limiar, são portadores de uma combinação de poligenes deletérios, somada à influência ambiental adversa, cuja ação conjunta não contribuiu para o desenvolvimento fenotípico normal. Esses indivíduos apresentam fatores genéticos que os tornam mais suscetíveis a manifestarem certos distúrbios em relação à população geral.

Os exemplos de fenótipos multifatoriais com limiar englobam inúmeras doenças humanas comuns na população, com nítida agregação familiar, como diabetes, asma, esquizofrenia, epilepsia, autismo, obesidade, doenças cardiovasculares, transtorno bipolar, esclerose múltipla, doença de Alzheimer, doença de Parkinson e malformações congênitas (anencefalia, espinha bífida, lábio e/ou palato fendido, hidrocefalia, microcefalia, onfalocele, pé torto, luxação congênita do quadril, estenose pilórica e cardiopatias congênitas).

Figura 6.4 Curva normal que mostra a distribuição de suscetibilidade para um fenótipo quantitativo dicotômico hipotético em uma população. O limiar distingue duas categorias de indivíduos: normais e afetados. A área sombreada evidencia a proporção de indivíduos que apresentam o fenótipo anormal, em função de terem ultrapassado o limiar crítico de propensão ao fenótipo.

Ser ou não afetado por algumas dessas condições a princípio se assemelha a outras doenças monogênicas simples, considerando que, em ambos os casos, nota-se um mesmo efeito fenotípico dicotômico, ou seja, ter ou não a característica. No entanto, embora os *traços poligênicos dicotômicos* possam ser confundidos com os *fenótipos qualitativos de origem monogênica*, é importante compreender que a presença da característica ocorre de maneira nitidamente diferente nas duas situações. Nos casos das doenças monogênicas, como vimos anteriormente, ter ou não determinado distúrbio ocorre em virtude de o indivíduo ser ou não portador de mutação em um único gene que leva ao fenótipo. Desse modo, indivíduos do sexo masculino portadores de uma alteração no gene *FMR1*, localizado no cromossomo X, que impeça sua ação e comprometa a síntese da proteína FMRP, terão a síndrome do X frágil, uma importante causa genética de deficiência mental. De outro modo, nas *doenças poligênicas com limiar*, a manifestação fenotípica dicotômica da condição se dá em decorrência da ação conjunta de poligenes que, em interação com agentes ambientais adversos, predispõem ao aparecimento da condição. A combinação desses fatores dará ao indivíduo um risco maior ou menor de desenvolver a doença, sendo que só manifestarão a condição aqueles cuja combinação de fatores de risco ultrapassar o limiar crítico de propensão ao fenótipo.

Variância fenotípica total
Variabilidade fenotípica referente a um traço específico observado entre organismos de uma população

Variância genotípica
Porção da variabilidade fenotípica de um traço multifatorial decorrente das diferenças genéticas entre os organismos

Variância ambiental
Porção da variabilidade fenotípica de um traço multifatorial decorrente das variações ambientais

Herdabilidade

A *variabilidade fenotípica* notada em uma população em relação a qualquer traço quantitativo multifatorial é, como vimos, decorrente da ação de dois componentes: genético e ambiental. Sendo assim, podemos expressar a variabilidade de um traço, também denominada variância fenotípica total (σ^2_p), como a soma da variância genotípica (σ^2_g), que mede os efeitos das diferenças genéticas entre os organismos, com a variância ambiental (σ^2_e), que mede os efeitos das variações ambientais, como mostra a Equação 6.1.

$$\sigma^2_p = \sigma^2_g + \sigma^2_e \qquad \text{(Equação 6.1)}$$

Existe grande interesse em saber quanto da variação de um traço quantitativo, observada em uma população específica, é resultado de diferenças genotípicas entre os indivíduos e quanto dessa variabilidade fenotípica ocorre em virtude dos diferentes agentes ambientais aos quais são expostos. Busca-se intensamente ponderar o significado dos fatores genéticos em relação à influência do ambiente na variação fenotípica total, o que representa um grande desafio.

Contudo, como medir a contribuição da variância genética na variabilidade fenotípica total de uma condição multifatorial em uma população? Essa estimativa é possível por meio do

Herdabilidade
Mede a contribuição relativa da variância genética na variabilidade fenotípica total de uma condição multifatorial em uma população

Herdabilidade em sentido restrito (h^2)
Proporção da variância fenotípica total de um traço quantitativo decorrente da variância genética aditiva

Herdabilidade em sentido amplo (H^2)
Proporção da variância fenotípica total de um traço quantitativo que pode ser atribuída à variância genotípica total e não apenas aos efeitos aditivos dos poligenes

Variância genética aditiva
Fração da variância fenotípica total de uma característica multifatorial que decorre dos efeitos aditivos dos diferentes poligenes

cálculo da herdabilidade que, sendo uma relação entre variâncias, nos possibilita estimar a contribuição da variância genética na variância fenotípica total. Essa medida expressa, portanto, a proporção da variabilidade fenotípica de um traço quantitativo que pode ser atribuída aos genes em determinada população, em um ambiente específico, de maneira que quanto maior o valor da herdabilidade, maior será a contribuição das diferenças genéticas entre os indivíduos daquela população na variabilidade fenotípica observada. É importante destacar que estimativas da herdabilidade de uma característica específica podem variar entre populações.

A herdabilidade de um traço pode ser representada como a fração da variância fenotípica total que resulta da *variância genética aditiva* (σ^2_a/σ^2_p), constituindo o que denominamos herdabilidade em sentido restrito (h^2), ou ser expressa como a proporção da variância fenotípica total que é decorrente da *variância genotípica total* (σ^2_g/σ^2_p), denominada herdabilidade em sentido amplo (H^2).

O que distingue a variância genética aditiva (aquela que resulta da soma dos efeitos individuais dos vários alelos) da *variância genotípica total* é que esta, além dos efeitos aditivos dos poligenes, engloba também outros que podem contribuir para a variabilidade fenotípica, como as relações de dominância entre alelos e as interações epistáticas entre genes diferentes. Desse modo, o componente genético da Equação 6.1 poderia ser expresso como a soma da variância aditiva (σ^2_a), da variância de dominância (σ^2_d) e da variância da interação ou epistática (σ^2_i), sendo essa equação representada por:

$$\sigma^2_p = \sigma^2_e + \underbrace{\sigma^2_g}_{(\sigma^2_a + \sigma^2_d + \sigma^2_i)} \qquad \text{(Equação 6.2)}$$

O valor da herdabilidade pode variar de 0 a 1 e, quanto mais próximo de 1, maior será a fração da variabilidade de um traço multifatorial que pode ser atribuída às diferenças genéticas, o que significa que nossa capacidade de prever o fenótipo da prole será maior. Valores que tendem a 1 indicam, portanto, que grande parte da variância fenotípica observada em uma população pode ser atribuída à variação genotípica, sendo o componente ambiental de menor significado na variação fenotípica total. Em contraposição, valores próximos a 0 indicam que as diferenças genéticas entre os indivíduos pouco ou nada contribuem para a variância fenotípica observada. A Tabela 6.2 apresenta estimativas da herdabilidade para algumas condições multifatoriais com limiar. Observe que a herdabilidade da asma foi estimada em 0,8, o que significa que cerca de 80% da variação total desse fenótipo pode ser atribuída a diferenças genotípicas entre os indivíduos da população estudada e 20% a influências ambientais.

Tabela 6.2 Estimativas da herdabilidade para distúrbios multifatoriais dicotômicos.

Distúrbio	Herdabilidade
Esquizofrenia	0,85
Asma	0,8
Fenda labial com ou sem fenda palatina	0,76
Estenose pilórica	0,75
Pé torto	0,68
Luxação congênita do quadril	0,60
Defeitos de fechamento do tubo neural (anencefalia e espinha bífida)	0,60

Método de gêmeos
Refere-se ao estudo comparativo entre pares de gêmeos monozigóticos e dizigóticos, bastante utilizado na estimativa da herdabilidade

Gêmeos monozigóticos
São aqueles originados da divisão e separação de um único zigoto e, por essa razão, possuem genótipos idênticos

Estudos de gêmeos

É possível estimar a *herdabilidade* de um traço multifatorial por meio da aplicação de diferentes métodos estatísticos. Em humanos, esse cálculo pode ser feito de correlações fenotípicas entre parentes próximos, como pais e filhos ou irmãos, lembrando que parentes próximos, além dos genes, em geral também compartilham os mesmos ambientes. Outro modo de calcular a *herdabilidade* é pelos *estudos comparativos de gêmeos* (método de gêmeos), bastante utilizados na estimativa da *herdabilidade* de traços quantitativos com limiar. Por meio desse método podemos comparar a manifestação de um fenótipo específico entre gêmeos monozigóticos (MZ),

Gêmeos dizigóticos
São aqueles originados de uma ovulação materna dupla e de dois eventos independentes de fertilização, sendo geneticamente tão semelhantes quanto dois irmãos não gêmeos

também denominados idênticos ou univitelinos, e entre gêmeos dizigóticos (DZ), também referidos como *fraternos* ou *bivitelinos*. Os gêmeos MZ, originados a partir da divisão e separação de um único zigoto, são geneticamente iguais, embora possam apresentar diferenças na expressão de alguns genes. De modo oposto, os gêmeos DZ, originados da ovulação e fecundação dupla, representam genomas distintos, sendo geneticamente tão semelhantes quanto dois irmãos não gêmeos. A incidência de gêmeos MZ é similar nas diversas populações humanas e representa cerca de 0,3% dos nascimentos, enquanto a frequência de gemelaridade DZ varia entre 0,2% e 1% nos diferentes grupos étnicos, sem considerarmos os casos resultantes de tratamento hormonal para infertilidade.

A comparação das variâncias fenotípicas de um traço em pares de gêmeos MZ e DZ possibilita uma estimativa de H^2 (*herdabilidade em sentido amplo*). Entre pares de gêmeos MZ, a variância genotípica será 0, de maneira que as diferenças fenotípicas entre eles corresponderão à variância ambiental (σ^2_e). Por outro lado, a variabilidade fenotípica entre pares de gêmeos DZ ocorrerá em virtude da variância ambiental (σ^2_e) e da variância genotípica (σ^2_g).

Pelo *método de gêmeos*, pares MZ são comparados com gêmeos DZ de mesmo sexo e correlações são estabelecidas. Um par de gêmeos (MZ ou DZ) pode ser concordante ou discordante em relação a determinado traço e os membros de um par serão concordantes quando ambos ou nenhum deles manifestarem a característica. Se um expressar o traço e o outro não, o par de gêmeos será discordante. Ao analisar a manifestação de um traço dicotômico, como a diabetes, em 100 pares de gêmeos, verificamos que 75 pares eram concordantes (ambos eram afetados ou não pela doença), o que nos leva a concluir que a taxa de concordância é de 75/100 (75%), o que significa que 25% dos pares de gêmeos apresentaram discordância em relação à manifestação da doença. Essa estimativa, embora não seja definitiva, é de grande valor, pois, das comparações das *taxas de concordância entre gêmeos MZ e DZ* podemos estimar a herdabilidade de vários traços multifatoriais.

É importante notar que pares de gêmeos, em geral, compartilham influências ambientais muito semelhantes, de maneira que quanto maior for a taxa de concordância de um traço entre pares de gêmeos MZ, em comparação com DZ, maior serão os indícios de que um componente genético importante seja responsável pela variabilidade fenotípica observada, pois seria esperado que gêmeos geneticamente idênticos apresentassem uma concordância maior dos que os DZ, geneticamente relacionados, porém não idênticos. Naqueles fenótipos nos quais se observa uma alta taxa de concordância tanto entre pares de gêmeos MZ quanto DZ, pode-se prever uma contribuição expressiva do ambiente. Fenótipos que se apresentam semelhantes em ambientes distintos provavelmente estão associados a um forte componente genético.

A concordância maior de um traço entre gêmeos MZ em relação aos DZ é um indicador da significância dos fatores genéticos na variabilidade do traço.

Loci de características quantitativas (QTL)

Embora o cálculo da herdabilidade nos forneça estimativa da influência genética na variabilidade fenotípica de um traço multifatorial, ela nada revela sobre quantos e quais são os *loci* e alelos que contribuem para a variação fenotípica observada, se eles influem igualmente ou de maneira diferenciada no fenótipo ou, mesmo, como ocorre a interação do ambiente com os poligenes. Um gene que contribui para um traço quantitativo é denominado *locus de característica quantitativa* (QTL, *quantitative trait locus*). De maneira conjunta, os múltiplos genes que contribuem para um fenótipo poligênico são os *loci de característica quantitativa* (QTL).

Apesar de nosso entendimento sobre QTL relacionados com fenótipos poligênicos humanos ainda ser muito restrito, na última década, conhecimentos sobre a estrutura do genoma humano têm possibilitado a construção de mapas densos de marcadores genéticos polimórficos, os quais têm servido de base para alcançarmos informações importantes relativas à base genética de traços multifatoriais, particularmente no que se refere à identificação de fatores genéticos que contribuem para o desenvolvimento de doenças complexas. Sabemos que as sequências do genoma humano têm, aproximadamente, 99,9% de semelhança entre indivíduos, independentemente de sua origem étnica, e que grande parte da variabilidade genética presente em nosso genoma ocorre em forma de variações em bases únicas do DNA, denominadas

G·F Genética em Foco

QTL | Bancos de dados de animais e plantas

Recentes conhecimentos na área da genômica, aliados a avanços tecnológicos que possibilitam a varredura em larga escala de todo o genoma, têm possibilitado a identificação de QTL, promovendo progressos expressivos no estudo de variação genética quantitativa, o que nos possibilitará responder de modo satisfatório às questões em relação a alguns traços. Na busca por genes candidatos, o mapeamento de QTL tem sido bem-sucedido em uma ampla variedade de organismos-modelo, animais e vegetais, o que beneficia diversos setores, como a agricultura e a criação de animais. Encontram-se disponíveis, para consulta na internet, bancos de dados de QTL de animais (*Animal QTL database*) e de plantas (*Gramene QTL database*). O *Animal QTL database* inclui mais de 15.000 QTL relacionados com 1.418 caracteres quantitativos no porco, gado, galinha e ovelha. O *Gramene QTL database* contém dados de espécies diferentes de plantas, incluindo arroz, milho, trigo, aveia e cevada, e registra mais de 11.500 QTL.

Polimorfismo de nucleotídio único (SNP, do inglês *single nucleotide polymorphism*) Variação de uma única base do DNA observada no genoma de indivíduos de uma população

polimorfismos de nucleotídios únicos (SNP, do inglês *single nucleotide polymorphisms*). Os SNP são responsáveis não apenas por traços relacionados à nossa fisiologia, psicologia e comportamento, mas também pela suscetibilidade a diversos distúrbios, já que podem influenciar a manifestação de determinadas doenças. Os bancos de dados (Ensembl v.63) contabilizam 30.095.750 SNP espalhados no genoma nuclear humano e essa abundância de variantes representa uma das mais significativas fontes de variabilidade genética nas populações, o que os torna excelentes marcadores da variação genética individual.

Chip de DNA Arranjo ordenado de sequências de DNA fixadas sobre um microssuporte sólido usado para análises de ligação com uma sequência-teste (DNA ou RNA)

Entre as diferentes estratégias utilizadas na identificação de QTL, destacamos os *estudos de associação envolvendo todo o genoma* (GWAS, *genome wide association studies*), um poderoso método capaz de rastrear variações no genoma inteiro pela genotipagem simultânea de mais de 500.000 SNP em um único *microarranjo de DNA* (*chip* de DNA). A aplicação dessa metodologia possibilita definir se determinadas variantes genéticas ocorrem com mais frequência em pessoas com um fenótipo clínico específico do que na população em geral, de maneira que a comparação de grandes amostras de pacientes e indivíduos controles normais nos possibilita fazer associações entre variantes genéticas e a predisposição a determinadas doenças. De fato, com base em *estudos de associação genômica*, tem sido possível correlacionar a presença de diversos SNP com o risco de desenvolvimento de distúrbios comuns na população, como:

G·F Genética em Foco

Doença de Alzheimer | Novas evidências de GWAS

Recente meta-análise de associação genômica, relativa à doença de Alzheimer, possibilitou a identificação de dois novos *loci* para essa doença neurodegenerativa. Na pesquisa, foram avaliados 529.205 SNP em pouco mais de 16.000 indivíduos, dos quais 5.964 tinham a doença de Alzheimer. Além de confirmar a existência de uma associação de risco entre essa doença e alelos do gene *APOE*, o estudo revelou uma nova associação com 2 SNP localizados em um íntron do gene *CLU* e na região 5′ do gene *PICALM*. A pesquisa mostrou que, além dos polimorfismos do gene *APOE*, variantes genéticas em *CLU* e *PICALM* também constituem fatores de suscetibilidade para a doença de Alzheimer.

doenças cardiovasculares e coronarianas, asma, diabetes, transtorno bipolar, artrite reumatoide, doença de Alzheimer e doença de Parkinson.

Apesar de as descobertas derivadas dos GWAS ainda serem modestas, elas representam um passo importante na direção da identificação de variantes genéticas relacionadas a doenças humanas complexas, pois possibilitam que as variantes genômicas encontradas sejam sistematicamente testadas em relação a sua contribuição na predisposição a esses distúrbios em diferentes populações. Isso poderá resultar em melhor compreensão dos processos biológicos subjacentes a essas patologias, abrindo caminhos para o delineamento de novas estratégias preventivas, diagnósticas e terapêuticas para estas enfermidades. Podemos antever que, à medida que nossos conhecimentos sobre fatores genéticos de risco avancem, testes genéticos poderão ser disponibilizados, tornando possível a identificação dos indivíduos geneticamente suscetíveis à determinada doença complexa.

É importante destacar, também, que a ênfase dada aos estudos de associação genômica, na busca de alcançar mais conhecimento sobre os fatores genéticos de predisposição relacionados com doenças multifatoriais, não deve obscurecer a investigação dos fatores ambientais (protetores e nocivos) nessas condições. É notório o efeito altamente benéfico da suplementação do ácido fólico na dieta como maneira de evitar defeitos no fechamento do tubo neural, que resultam em anencefalia ou espinha bífida. A doença de Parkinson é também um bom exemplo de distúrbio complexo, no qual alguns fatores ambientais protetores (consumo de cafeína) e prejudiciais (exposição a pesticidas) são conhecidos. Entretanto, os mecanismos de interação entre poligenes e ambiente que constituem a base dessas doenças ainda são, em sua grande maioria, desconhecidos.

RESUMO

- A maior parte da variabilidade fenotípica observada nos organismos vivos é de natureza poligênica, ou seja, resulta da ação combinada de alguns ou de vários genes em diferentes *loci* (poligenes), podendo também sofrer a influência de fatores ambientais, constituindo a herança multifatorial
- A variabilidade fenotípica dos traços quantitativos resulta da combinação entre genes diferentes e de complexas interações alélicas intralocus e entre *loci*, de maneira que a contribuição dos poligenes para determinada característica quantitativa possa ocorrer de maneira desigual
- A variação fenotípica de origem monogênica é de natureza qualitativa e resulta da expressão de um gene único que produz um efeito nítido no fenótipo. De maneira oposta, os fenótipos poligênicos são de natureza quantitativa e exibem variação contínua nas populações, ou seja, se distribuem gradativamente entre dois extremos
- A análise do heredograma não contribui para as previsões sobre a transmissão dos fenótipos quantitativos, sendo necessário dispor de dados populacionais dos quais inferências estatísticas possam ser derivadas
- À medida que o número de genes relacionados com um traço quantitativo contínuo aumenta, verifica-se também um número crescente de classes fenotípicas, o que resulta em distribuição cada vez mais uniforme, na forma de uma curva normal ou gaussiana
- Na distribuição normal, a maioria dos fenótipos quantitativos se distribui ao redor da média, sendo que 68%, 95% e 99,7% dos valores de uma variável encontram-se no intervalo delimitado pela média mais ou menos um, dois ou três desvios padrão, respectivamente
- As características quantitativas podem ser separadas em duas categorias: poligênicas contínuas e poligênicas dicotômicas
- As características poligênicas contínuas são representadas por variáveis que podem ser quantificadas ou medidas e que exibem variação contínua na população, sem uma separação nítida de categorias
- Outras características poligênicas se apresentam de maneira dicotômica na população. Esse tipo de variação fenotípica pode ser compreendido do modelo limiar ou de propensão, o qual estabelece a existência de um limiar crítico na curva de distribuição normal, que distingue duas categorias fenotípicas nitidamente opostas
- Quanto maior a influência do componente genético em relação aos fatores ambientais na variância fenotípica de um traço, mais genético esse traço será e quanto menor a influência do genótipo sobre a variabilidade fenotípica, menos genético ele será
- Todos os organismos apresentam uma plasticidade fenotípica, ou seja, são sensíveis às variações ambientais e capazes de expressar fenótipos diferentes em ambientes diversos
- A variabilidade fenotípica relativa a qualquer traço multifatorial pode ser representada como a soma da variância genotípica, que mede os efeitos das diferenças genéticas entre os organismos, com a variância ambiental, que mede os efeitos das variações no ambiente
- A contribuição da variância genética na variabilidade fenotípica total de um traço multifatorial pode ser estimada pelo cálculo da herdabilidade, que varia de 0 a 1. Quanto maior o valor da herdabilidade, maior será a contribuição das diferenças genéticas entre os indivíduos daquela população na variabilidade fenotípica observada
- A herdabilidade pode ser representada como a fração da variância fenotípica total que resulta da variância genética aditiva (herdabilidade em sentido restrito) ou ser expressa como a

proporção da variância fenotípica total que é devida à variância genotípica total (herdabilidade em sentido amplo)

- Um modo de calcular a herdabilidade para traços multifatoriais humanos é pelo estudo comparativo de gêmeos monozigóticos e dizigóticos (método de gêmeos), que utilizam taxas de concordância entre pares de gêmeos. A concordância maior de um traço entre gêmeos MZ em relação aos DZ é um indicador da significância dos fatores genéticos na variabilidade do traço
- Um gene que contribui para um traço quantitativo é denominado *locus* de característica quantitativa (QTL, *quantitative trait locus*) e a identificação de QTL tem sido possível em decorrência dos novos conhecimentos e avanços tecnológicos na área da genômica
- Pelos estudos de associação genômica, inúmeros polimorfismos de nucleotídios únicos (SNP), que representam uma importante classe de marcadores da variação genética em humanos, têm sido identificados como fatores de risco para o desenvolvimento de doenças multifatoriais nas populações.

AUTOAVALIAÇÃO

6.1 Diferencie herança monogênica de herança poligênica e multifatorial.
6.2 Em relação à variabilidade fenotípica presente nas populações, qual é a base genética da variação qualitativa e quantitativa?
6.3 Que são poligenes e que tipos de interações possíveis entre eles contribuem para a variação genética quantitativa?
6.4 A análise de heredograma contribui para estabelecermos previsões sobre a transmissão de fenótipos quantitativos? Explique.
6.5 O que diferencia a herança poligênica contínua da herança poligênica dicotômica?
6.6 Explique de que modo os agentes ambientais contribuem para a variação fenotípica.
6.7 Defina plasticidade fenotípica e a norma de reação.
6.8 Qual é o conceito de fenótipo quantitativo contínuo, categórico e dicotômico?
6.9 De que modo as características quantitativas se distribuem nas populações e quais são os parâmetros estatísticos que podem ser utilizados na avaliação dessa distribuição?
6.10 Comente: "Embora os traços poligênicos dicotômicos possam ser confundidos com os fenótipos qualitativos monogênicos, é importante compreender que a presença do traço ocorre de maneira nitidamente distinta nas duas situações."
6.11 Explique a variação fenotípica quantitativa dicotômica à luz do modelo limiar.
6.12 Quais são os componentes da variância fenotípica?
6.13 Em que consiste a variância genética aditiva?
6.14 Conceitue herdabilidade, diferenciando herdabilidade em sentido amplo de herdabilidade em sentido restrito.
6.15 O que significa dizer que a herdabilidade de determinado traço quantitativo é de 0,7?
6.16 O que é o método de gêmeos? Qual é a sua aplicabilidade?
6.17 Defina QTL.
6.18 De que modo os estudos de associação genômica podem contribuir para a identificação de QTL?
6.19 Faça uma consulta na internet aos bancos de dados de QTL de animais e de plantas, buscando atualizar os dados sobre poligenes das diferentes espécies.
6.20 Por que os SNP são considerados importantes marcadores da variação genética individual e qual é a relação deles com as doenças humanas complexas?
6.21 O que você entende por *chip* de DNA?

BIBLIOGRAFIA

ANIMAL QTL DATABASE. Disponível em: http://www.animalgenome.org/cgi-bin/QTLdb/index. Acessado em 22/02/2012.

BERTRAM, L.; LANGE, C.; MULLIN, K. *et al.* Genome-wide association analysis reveals putative Alzheimer's disease susceptibility *loci* in addition to APOE. *Am J Hum Genet* 2008; 83: 623-632.

BOTSTEIN, D.; RISCH, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease *Nature Genetics Suppl* 2003; 33: 228-237.

ENSEMBL v.63 (June 2011). Disponível em: http://www.ensembl.org/index.html. Acesso em 22/02/2012.

GRAMENE QTL DATABASE http://www.gramene.org/qtl/. Acesso em 22/02/2012.

GRIFFITHS, A.J.F.; WESSLER, S.R.; LEWONTIN, R.C. *et al. Introdução à genética.* 9ª ed. Rio de Janeiro: Guanabara Koogan, 2008.

HAROLD, D.; ABRAHAM, R.; HOLLINGWORTH, P. *et al.* Genome-wide association study identifies variants at *CLU* and *PICALM* associated with Alzheimer's disease. *Nat Genet* 2009; 41: 1088-1093.

HARTL, D.L.; CLARK, A.G. *Princípios de genética de populações.* 4ª ed. Porto Alegre: Artmed, 2010.

HAWKES, C. H. Twin studies in medicine – what do they tell us? *Q J Med* 1997; 90: 311-321.

KAJIWARA, K.; BERSON, E.L.; DRYJA, T.P. Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/*RDS* and *ROM1 loci. Science* 1994; 264: 1604-1608.

KLUG, W.S.; CUMMINGS, M.R.; SPENCER, C.A.; PALLADINO, M.A. *Conceitos de genética.* 9ª ed., Porto Alegre: Artmed, 2010.

LAMBERT, J.C.; HEATH, S.; EVEN, G. *et al.*. Genome-wide association study identifies variants at *CLU* and *CR1* associated with Alzheimer's disease. *Nat Genet* 2009; 41: 1094-1099.

LI, H.; WETTEN, S.; LI, L. *et al.* Candidate single-nucleotide polymorphisms from a genome-wide association study of Alzheimer disease. *Arch Neurol* 2008; 65: 45-53.

MACCARROLL, S. A. Extending genome-wide association studies to copy-number variation. *Hum Mol Genet* 2008; 17: 135-142.

MACKAY, T.F.C. The genetic architecture of quantitative traits. *Ann Rev Genet* 2001; 35: 303-339.

MITCHELL, L. E. Epidemiology of neural tube defects. *Am J Med Genet C Semin Med Genet* 2005;135: 88-94.

MOFFATT, M.F.; KABESCH, M.; LIANG, L. *et al.* Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; 448: 470-473.

MOORE, J. H.; WILLIAMS, S. M. Epistasis and its implications for personal genetics. *Am J Hum Genet* 2009; 85: 309-320.

NUSSBAUM, R.L.; MCINNES, R.R.;WILLARD, H.F. *Thompson & Thompson. Genética médica.* 7ª ed., Rio de Janeiro: Elsevier, 2007.

PASTERNAK, J.J. *Uma introdução à genética molecular humana.* 2ª ed., Rio de Janeiro: Guanabara Koogan, 2007.

PRICE, A. H. Believe it or not, QTLs are accurate! *Trends Plant Sci* 2006; 11: 213-216.

SATAKE, W.; NAKABAYASHI, Y.; MIZUTA, I. *et al.* Genome-wide association study identifies common variants at four *loci* as genetic risk factors for Parkinson's disease. *Nat Genet* 2009; 41: 1303-1307.

SAXENA, R.; VOIGHT, B.F.; LYSSENKO, V. *et al.* Genome-wide association analysis identifies *loci* for type 2 diabetes and triglyceride levels. *Science* 2007; 316: 1331-1336.

SIMÓN-SÁNCHEZ, J.; SCHULTE, C.; BRÁS, J. M. Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet* 2009; 41: 1308-1314.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

TANNER, C.M. Advances in environmental epidemiology. *Movement disorders* 2010; 25, Suppl. 1: S58–S62.

TURNPENNY, P.D.; ELLARD, S. *Emery genética médica.* 13ª ed., Rio de Janeiro: Elsevier, 2009.

ZEGGINI, E.; SCOTT, L.J.; SAXENA, R. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility *loci* for type 2 diabetes. *Nat Genet* 2008; 40: 638-645.

7

Genética Metabólica, Hemoglobinopatias e Imunogenética

Objetivos de estudo

Compreender três importantes processos de manutenção do meio interno constante em humanos: as principais vias metabólicas, o transporte de oxigênio e a resposta imunológica

Conhecer as principais vias metabólicas relacionadas com erros inatos do metabolismo

Definir e avaliar a importância das mutações gênicas que afetam as principais vias metabólicas e patologias associadas

Conhecer os principais aspectos do transporte de oxigênio em humanos

Definir e avaliar a importância dos diferentes tipos de hemoglobina causados por mutações gênicas e patologias associadas

Conhecer os principais processos de reconhecimento do que é próprio e não próprio em humanos

Definir e avaliar a importância dos mecanismos genéticos envolvidos na resposta imunológica

Conceitos-chave do capítulo

Alergia
Aminoácido essencial
Antígenos
ATP
Carboidratos
Células auxiliares de memória
Células citotóxicas naturais
Células fagocitárias
Choque anafilático
Citocinas
Complexo principal de histocompatibilidade ou MHC
Doenças autoimunes
Efeito do fundador
Endocitose
Endossomos
Enzimas
Erros inatos do metabolismo
Estrutura quaternária
Família gênica
Genômica
Heme
Hemoglobina
Hemoglobinopatias
Imunização ativa
Imunização passiva
Imunoglobulinas
Interferons
LDL
Linfócitos B
Linfócitos de memória
Linfócitos T
Linfócitos T auxiliares
Lipoproteínas
Lisossomos
Metabolismo
Mutação gênica
Porfirina
Pseudogenes
Receptores celulares
Sistema imunológico
Substâncias farmacologicamente ativas
Sistema imunológico
Sistema complemento
Terapia gênica
Vacina

Introdução

Genômica
Ramo da genética que descreve e analisa a constituição do genoma de uma espécie. Seu desenvolvimento é muito recente e depende de genomas completamente sequenciados e de ferramentas de bioinformática

As doenças genéticas que envolvem as vias de manutenção do meio interno constante estão entre as mais estudadas pelos pesquisadores. No entanto, elas continuam a ser consideradas um grande desafio para a ciência. Os últimos avanços em genômica humana, além de mostrar que os seres humanos têm enorme variabilidade genética, têm indicado um número cada vez maior de alterações patogênicas nos genes, associadas a distúrbios de vias importantes na fisiologia normal humana. Assim, neste capítulo iremos abordar os aspectos genéticos ligados ao metabolismo, ao transporte de oxigênio através das hemoglobinopatias e ao sistema imunológico.

Doenças genéticas do metabolismo

Mutação gênica
Alteração na sequência de nucleotídios do DNA que constitui o gene

As doenças genéticas do metabolismo se caracterizam pela participação de mutações gênicas reconhecidamente patogênicas, ou seja, mutações que vão alterar o produto gênico de modo a determinar o desenvolvimento de doenças, impedindo ou prejudicando vias participantes do metabolismo normal da célula. Essa parte da genética também é conhecida como genética bioquímica.

Assim, as doenças genéticas do metabolismo constituem a genética bioquímica que, por sua vez, descreve as doenças de caráter hereditário que envolvem vias do metabolismo celular, promovendo distúrbios bioquímicos no organismo.

O termo erros inatos do metabolismo foi cunhado pelo médico inglês Archibald Garrod, há mais de 100 anos, para designar o conjunto de doenças metabólicas de caráter hereditário. Esse pesquisador foi pioneiro na descrição desse tipo de doença, demonstrando uma série de distúrbios bioquímicos que se apresentavam caracteristicamente familiares. Ele observou que uma enzima funcionalmente alterada prejudica uma via metabólica, levando ao desenvolvimento da doença. Dessas observações, mais o desenvolvimento de conceitos como um gene – uma enzima (ver Genética em Foco), determinou-se que a maioria das doenças genéticas bioquímicas pode ser explicada pela deficiência de uma enzima atuante em via metabólica por mutação gênica, resultando em prejuízo para o bom funcionamento do organismo.

Erros inatos do metabolismo
Termo utilizado para descrever o conjunto de doenças genéticas do metabolismo

Enzimas
Proteínas ou RNA com características catalíticas, ou seja, que promovem reações químicas em sistemas vivos

Metabolismo
Conjunto de processos físico-químicos que transformam substâncias dentro de um organismo. As reações bioquímicas que fazem parte do metabolismo estão organizadas em sequência, chamadas de vias metabólicas, nas quais o produto de uma reação é utilizado como substrato para a reação seguinte

G·F Genética em Foco

Atualização do conceito: um gene – uma enzima

Um conceito importante para a compreensão das doenças genéticas do metabolismo foi a ideia de que um gene é responsável pela produção de uma enzima. O conceito um gene – uma enzima foi desenvolvido pelos pesquisadores Beadle e Tatum (hipótese Beadle e Tatum), pela realização de experimentos com o fungo *Neurospora crassa*, pelos quais essa dupla de pesquisadores recebeu o prêmio Nobel de Medicina de 1958. Esses experimentos foram feitos pela irradiação das células desse fungo com raios X, estabelecendo quais enzimas e, portanto, quais vias metabólicas estavam afetadas pela presença de mutações. Beadle e Tatum, junto com outros colaboradores, determinaram que um gene produz uma enzima que é responsável por um passo metabólico. Eles descreveram, então, uma relação direta entre a mutação em um gene e um passo da via afetada. Esse conceito facilitou o entendimento das doenças genéticas metabólicas. A evolução desse conceito levou à formulação da afirmação "um gene-um polipeptídio", já que diversos estudos na época já descreviam produtos gênicos como polipeptídios, uma denominação geral para unidades proteicas como as que constituem, por exemplo, as hemoglobinas. Atualmente, com o grande avanço da genética molecular, sabemos que esses conceitos não são de todo corretos, já que parte dos nossos genes, individualmente, pode produzir diferentes proteínas pelo processamento alternativo de transcritos.

Como exemplo, temos a alcaptonúria, doença metabólica que originou as ideias do Dr. Garrod, e a deficiência da enzima oxidase do ácido homogentísico, na via de degradação da fenilalanina. Vários são os distúrbios bioquímicos causados por essa importante via do metabolismo de aminoácidos, sendo um importante exemplo para a compreensão das chamadas doenças genéticas do metabolismo. Hoje existe um grande número dessas doenças descritas e catalogadas em um banco de dados para doenças genéticas mendelianas, conhecido como MIM – Mendelian Inheritance in Man e sua versão *on-line* (OMIM®).

O OMIM® é uma importante fonte de informação *on-line* sobre genes humanos e fenótipos de doenças genéticas para médicos, geneticistas e estudantes da área de genética humana, sendo a primeira versão um catálogo de doenças genéticas de herança mendeliana.

Ele é atualizado diariamente e é uma marca registrada da Universidade Johns Hopkins, dos EUA. A maioria das doenças genéticas do metabolismo apresenta herança do tipo autossômica recessiva determinadas por herança monogênica. Essas doenças também podem ser do tipo autossômica dominante, ligadas ao X, recessivas ou dominantes, ou mitocondriais. A Tabela 7.1 apresenta as doenças genéticas do metabolismo, agrupadas segundo o tipo de herança e via afetada.

A seguir, abordaremos algumas delas.

■ **Tabela 7.1** Doenças genéticas do metabolismo segundo o seu tipo de herança.

Tipo de herança	Doença metabólica	Proteína afetada	OMIM
AD	Hipercolesterolemia familial	Receptor de LDL	143890
	Porfirias hepáticas		
	P. *variegata*	Protoporfirinogênio oxidase	176200
	P. intermitente aguda	Uroporfirinogênio I sintetase	176000
	Protoporfiria eritropoiética	Ferroquelatase	177000
AR	Metabolismo de aminoácidos		
	Fenilcetonúria	Fenilalanina hidroxilase	261600
	Alcaptonúria	Ácido homogentísico oxidase	203500
	Albinismo oculocutâneo	Tirosinase	203100
	Ciclo da ureia		
	Citrulinemia	Ácido argininosuccínico sintase	215700
	Hiperargininemia	Arginase	207800
	Metabolismo de carboidratos		
	Galactosemia	Gal-1-P uridiltransferase	230800
	Intolerância congênita à frutose	Aldolase B	229600
	Doenças de armazenamento do glicogênio		
	Doença de von Gierke (GSD-Ia)	Glicose-6-fosfatase	232200
	Doença de Pompe (GSD-II)	Alfa-1,4-glicosidase lisossômica	232300
	Doença de Cori (GSD-III)	Amilo-1,6-glicosidase	610860
	Doença de Andersen (GSD-IV)	Enzima desramificadora de glicogênio	232500
	Doença de McArdle (GSD-V)	Fosforilase muscular	232600
	Doenças de armazenamento lisossômico		
	Mucopolissacaridoses		
	Síndrome de Hurler (MPS-I)	Alfa-1-iduronidase	607014
	Síndrome de Sanfilippo (MPS-IIIa)	Heparan-S-sulfaminidase	252900
	Síndrome de Sly (MSP-VII)	Betaglicuronidase	253220
	Síndrome de Maroteaux-Lamy	Alfa-4-S-sulfatase	253200
	Esfingolipidoses		
	Doença de Tay-Sachs	Hexosaminidase-A	272800
	Doença de Gaucher, tipo 1	Glucocerebrosidase	231000
	Doença de Niemann-Pick, tipo A	Esfingomielinase	257200
	Metabolismo de esteroides		
	Hiperplasia adrenal congênita	21-hidroxilase	201910
	Metabolismo de purinas/pirimidinas		
	Acidúria orótica hereditária	Orotidina-5-P-descarboxilase	258900
	Imunodeficiência congênita grave causada por deficiência de ADA	Adenosina desaminase	102700
	Distúrbios peroxissômicos		
	Síndrome de Zellweger	Proteínas de biossíntese do peroxissomo	214100
	Doença de Wilson	ATPase transportadora de cobre	277900
LX	Insensibilidade androgênica	Receptor de andrógenos	300068
	Síndrome de Hunter (MPS-II)	Iduronato-S-sulfatase	309900
	Síndrome de Lesch-Nyhan	Hipoxantina guanina fosforibosiltranferase	300322
	Doença de Menkes	ATPase transportadora de cobre	309400
	Deficiência de ornitina transcarbamilase	Ornitina transcarbamilase	311250
Mit	Doença de Leigh	Complexos mitocondriais I, II, III, IV e V	256000
	Síndrome de Kearns-Sayre	Proteínas mitocondriais diversas	530000
	Síndrome de Pearson	Proteínas mitocondriais diversas	557000
	Neuropatia óptica hereditária de Leber	Proteínas mitocondriais diversas	535000

AD – autossômica dominante; AR – autossômica recessiva; LX – ligada ao cromossomo X; MIT – mitocondrial.

Doenças genéticas ligadas ao metabolismo de fenilalanina e tirosina

Fenilalanina e tirosina são aminoácidos que compõem nossas proteínas. Considera-se a fenilalanina um aminoácido essencial, ou seja, não pode ser sintetizado por seres humanos e tem de estar presente em nossa dieta. A tirosina, por sua vez, é sintetizada com base na fenilalanina, sendo o primeiro produto de sua via de degradação, a qual origina 2 produtos finais: o fumarato e o acetoacetato. Além de participar da formação das proteínas, a tirosina é fundamental para o organismo humano, já que é desse aminoácido que se formam importantes

Aminoácido essencial
Tipo de aminoácido que tem de estar presente na dieta humana, uma vez que não existem vias para sua biossíntese. Caso esse tipo de aminoácido não seja fornecido, todas as proteínas que apresentam o aminoácido em sua estrutura não serão sintetizadas

Substâncias farmacologicamente ativas
Tipo de substâncias ou moléculas orgânicas presentes nos organismos que funcionam modulando e alterando a atividade celular

substâncias farmacologicamente ativas, atuantes no controle do metabolismo, como os hormônios da tireoide tri-iodotirosina e tiroxina, as catecolaminas dopamina, norepinefrina e epinefrina e, ainda, a não menos importante melanina. Portanto, as vias que fazem parte da metabolização da fenilalanina e da tirosina são necessárias para a manutenção e o bom funcionamento do organismo como um todo.

Como anteriormente citado, a alcaptonúria, quando descrita, deu origem à investigação de uma nova classe de doenças – as doenças genéticas do metabolismo. A Figura 7.1 mostra as vias e as doenças metabólicas citadas associadas à disfunção das enzimas.

A fenilcetonúria (PKU), dita clássica, é a causa mais comum de hiperfenilalaninemias, que é o excesso de fenilalanina no organismo. Nesse caso, esse excesso é resultante do impedimento da atividade da enzima fenilalanina hidroxilase por mutação gênica. A atividade da fenilalanina hidroxilase é converter a fenilalanina em tirosina sendo, portanto, esse passo metabólico responsável pelas doenças advindas da falta da tirosina. Essa doença causa retardo mental grave nas crianças não tratadas e, ainda, eczemas e cor clara de pele e olhos em virtude da falta da pigmentação. São observadas ainda alterações neurológicas causadas pelo acúmulo de fenilalanina. Essa doença teve importante papel no avanço das análises bioquímicas para recém-natos. Tais efeitos terríveis resultantes do excesso de fenilalanina podem ser evitados pela restrição desse aminoácido na dieta e com o suplemento de tirosina nas crianças afetadas. Assim, o conhecido Teste do Pezinho, que consiste na obtenção de gotas de sangue da picada no pé do recém-nascido, foi desenvolvido e é utilizado na detecção precoce dos indivíduos afetados pela PKU. Segundo o Ministério da Saúde (http://portal.saude.gov.br), o Programa Nacional de Triagem Neonatal (PNTN) prevê o diagnóstico de 4 doenças: hipotireoidismo congênito, fenilcetonúria, hemoglobinopatias e fibrose cística. Nesse programa, uma vez detectada a doença na triagem inicial, as famílias são orientadas a procurarem outros exames confirmatórios, visto que as doenças triadas no "teste do pezinho" são assintomáticas no período neonatal e que, portanto, não devem demorar em procurar a confirmação diagnóstica dos casos suspeitos. O risco é ocasionar sequelas graves e irreversíveis no desenvolvimento da criança, que só serão percebidas tardiamente.

Hemoglobinopatias
Conjunto de doenças resultantes de alterações patogênicas na proteína hemoglobina, presente nas hemácias, cuja principal função é transportar oxigênio dos pulmões para os tecidos

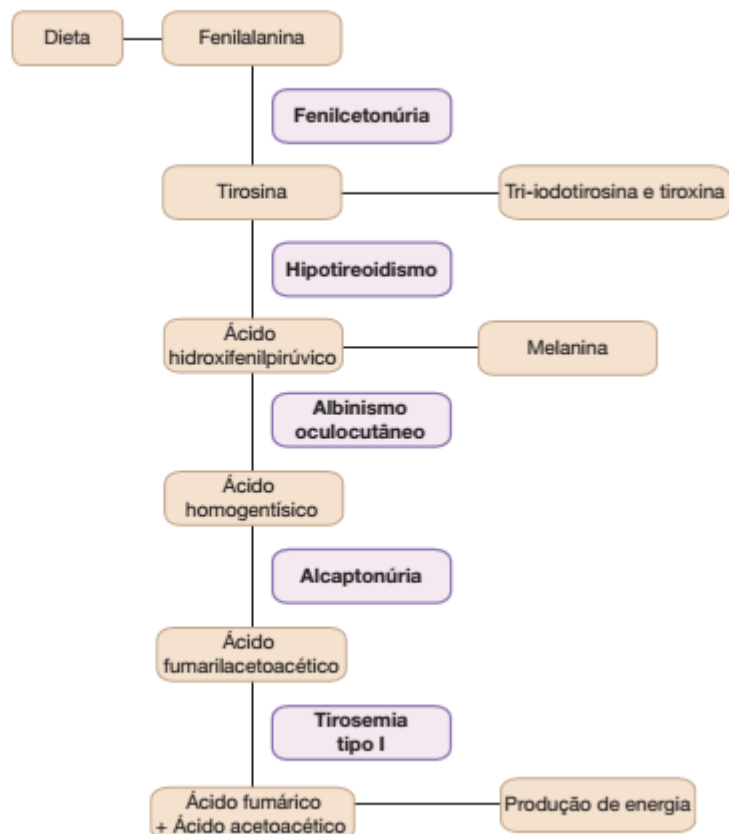



Figura 7.1 Vias e doenças metabólicas relacionadas com o metabolismo dos aminoácidos fenilalanina e tirosina.

A alcaptonúria já foi citada como a primeira doença genética identificada como erro inato do metabolismo. Os indivíduos portadores têm deficiência da enzima oxidase do ácido homogentísico e excretam quase toda a tirosina ingerida na urina, assim como produtos da oxidação do ácido homogentísico, o que dá a esta uma coloração escura característica.

O albinismo é a condição em que os indivíduos afetados não produzem melanina e, portanto, têm pele, cabelos, íris e fundo dos olhos despigmentados. Essa condição surge pela deficiência da enzima tirosinase, responsável pela síntese de melanina em várias etapas, transformando tirosina em eumelanina – tipo de melanina que corresponde ao pigmento escuro que funciona como substância fotoprotetora. A melanogênese é o processo de formação do pigmento melanina e acontece em células especializadas chamadas de melanócitos. O principal estímulo para a formação da melanina é a exposição dos melanócitos à luz ultravioleta do tipo B; portanto, os indivíduos portadores dessa doença apresentam sensibilidade à luz solar e podem desenvolver câncer de pele.

Doenças genéticas ligadas ao metabolismo de carboidratos: galactosemia e doença de armazenamento de glicogênio

Carboidratos
Classe de moléculas também conhecidas como glicídios ou açúcares, que podem ser classificados segundo quantidades de unidades de glicídios em poli-, tri-, di- ou monossacarídios

ATP
Da classe dos nucleotídios, a adenosina 5′-trifosfato (ATP) é uma molécula rica em energia, em virtude das ligações fosfato que, quando rompidas ou criadas, são utilizadas na transferência de energia química nas reações do metabolismo

O metabolismo de carboidratos é essencial para qualquer tipo de organismo, já que eles formam um grupo de compostos biológicos que, quando degradados, fornecem importante quantidade de energia para as células. A glicose é o melhor exemplo de carboidrato utilizado como fornecedor de energia, metabolizado pelo processo chamado de glicólise, que produz piruvato e adenosina trifosfato ou ATP. O piruvato é metabolizado pelo ciclo de Krebs que, acoplado à cadeia respiratória nos organismos aeróbicos, produz boa quantidade de energia sob a forma de ATP.

Vários carboidratos podem ser convertidos em glicose. Bons exemplos são o amido e glicogênio, polissacarídios cujas unidades formadoras são moléculas de glicose; a sacarose, um dissacarídio formado por frutose e glicose que constitui o açúcar comum, e os monossacarídios, frutose e galactose. Ambos são metabolizados no organismo por vias enzimáticas que, quando impedidas de seu funcionamento normal, leva a distúrbios metabólicos importantes. A galactosemia e a doença de von Gierke são exemplos de doenças genéticas ligadas ao metabolismo de carboidratos.

A galactosemia é uma condição patológica em que o indivíduo tem dificuldade em metabolizar a galactose.

A galactose, junto com uma molécula de glicose, constitui o dissacarídio lactose que é o principal açúcar do leite. Os recém-nascidos portadores dessa doença têm dificuldades em se nutrir, resultando em baixo peso, falta de energia e, ainda, danos no fígado. Crianças que não são tratadas apresentam dificuldade de desenvolvimento, com prejuízo na fala e no aprendizado e problemas nos olhos, como o aparecimento de catarata. Todos esses distúrbios são causados pelo excesso de metabólitos da galactose no organismo, o que pode ser evitado retirando-se leite e derivados da dieta. A galactosemia clássica acontece pela presença de mutações no gene que codifica a enzima galactose-1-fosfato uridil transferase, que participa da formação da glicose a partir da galactose. As doenças chamadas de armazenamento de glicogênio são várias e estão apresentadas na Tabela 7.1. Entre elas, a doença de von Gierke foi a primeira a ser descrita e é uma doença rara, também conhecida como doença de armazenamento de glicogênio tipo Ia. O glicogênio é um polissacarídio formado por inúmeras moléculas de glicose, fazendo o papel de estoque para essas moléculas. Esse polissacarídio está presente principalmente no fígado e nos músculos. Os problemas fisiológicos advindos da doença são resultantes da deficiência da enzima glicose-6-fosfatase, a qual catalisa a conversão da glicose-6-fosfato em glicose.

Os indivíduos portadores da doença de von Gierke, portanto, vão sofrer com a falta de glicose no organismo, sendo característica dessa doença o estado de hipoglicemia ou baixa de glicose no sangue.

Esses indivíduos apresentarão um quadro de fraqueza muscular, chegando até mesmo a resultar em insuficiência cardíaca. Como era de se esperar, o fígado é bastante afetado e normalmente se apresenta aumentado (hepatomegalia). O tratamento dos indivíduos é não

tornar possível o jejum, pela alimentação frequente, sempre com o objetivo de manter o nível adequado de glicose no sangue.

Sistema imunológico
Conjunto de reações do organismo que mantêm o equilíbrio interno constante, constituindo a defesa deste e eliminando agentes externos ou estranhos como bactérias e vírus. Ainda, o seu mau funcionamento pode levar ao desenvolvimento das doenças autoimunes e de hipersensibilidade

Doenças genéticas ligadas ao metabolismo de purinas e pirimidinas: Imunodeficiência congênita severa causada por deficiência de ADA

As purinas e pirimidinas são compostos nitrogenados fundamentais ao organismo. Eles fazem parte dos ácidos nucleicos (DNA e RNA) e de moléculas importantíssimas, como o ATP, fundamental na transferência de energia do metabolismo, já comentado anteriormente. Entre as doenças genéticas ligadas ao metabolismo de purinas e pirimidinas, a imunodeficiência congênita severa causada por deficiência de adenosina desaminase, ou ADA, se destaca. Essa doença afeta fundamentalmente o sistema imunológico do indivíduo portador, por uma deficiência imunológica severa, ligada ao impedimento da função normal de linfócitos B e T.

Linfócitos B e T
Células especializadas participantes da resposta imune dos organismos vertebrados

A imunodeficiência congênita severa causada por deficiência de ADA origina-se por mutação no gene da enzima ADA, mais precisamente a ADA1, responsável por uma das vias de degradação das purinas, as quais levam à formação do produto de excreção ácido úrico.

Apesar de não estar totalmente explicada a relação entre a deficiência da enzima e a baixa imunológica, os pesquisadores sugerem o importante papel da enzima para a formação e a manutenção dos ácidos nucleicos que, por sua vez, estão diretamente relacionados com a formação e à manutenção da função das células de defesa imunológica. Os linfócitos são continuamente formados em nosso organismo em virtude de estímulos. Portanto, a síntese de novas células fica prejudicada com a deficiência de ADA. Como sintomas da doença, crianças portadoras apresentam infecções recorrentes e, como o quadro clínico é bastante grave, elas podem vir a morrer no primeiro ano de vida. A imunodeficiência congênita severa causada por deficiência de ADA tem importância histórica, já que foi a primeira doença a ser tratada com sucesso por terapia gênica. Uma menina de 4 anos, com deficiência de ADA, em 1990, foi a primeira paciente a receber terapia gênica.

Terapia gênica
Tipo de terapia associado à introdução de genes em células ou indivíduos com o objetivo de corrigir o efeito patológico de genes mutantes

A Figura 7.2 mostra as etapas necessárias para o desenvolvimento de um tratamento por terapia gênica. Apesar do sucesso obtido nesse caso e em outros, esse tipo de terapia ainda encontra enormes desafios para ser de uso comum.

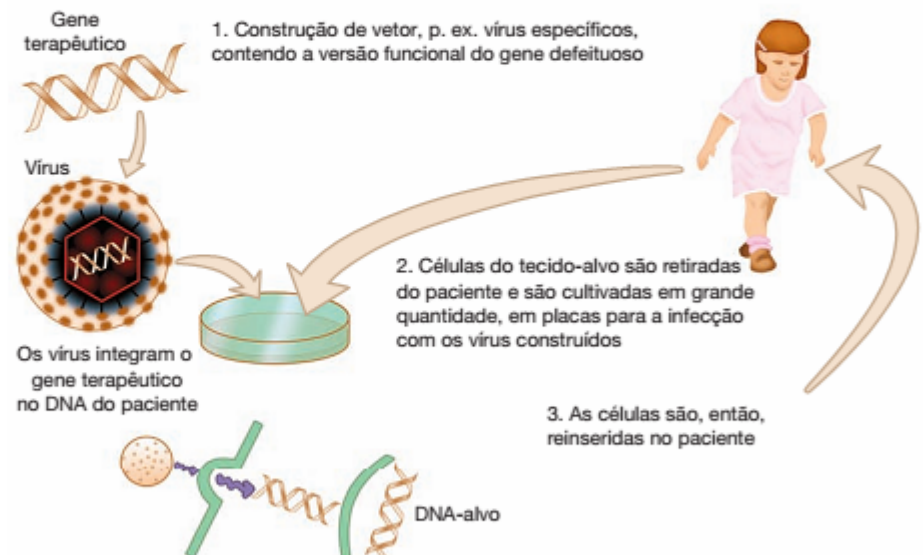



Figura 7.2 Principais etapas do desenvolvimento de terapia gênica em humanos.

Lipoproteínas
Classe de complexos moleculares formados por proteínas, chamadas de apolipoproteínas, e lipídios. Elas são identificadas segundo a densidade que depende do tipo de lipídio e têm, primariamente, a função de transporte para sua metabolização no fígado

LDL
Sigla que designa lipoproteína de baixa densidade. É a lipoproteína formada pela parte proteica, ou apolipoproteína, e moléculas de colesterol. As moléculas de colesterol dão o caráter de baixa densidade dessa lipoproteína

Endocitose
Processo pelo qual a célula pode internalizar substâncias ou organismos externos pela captação destas e envolvimento com a membrana celular

Endossomos
Organelas formadas pela ação da endocitose

Lisossomos
Organelas celulares envolvidas no processo de digestão de compostos intra e extracelulares

Heme
Grupamento químico constituído por um átomo de ferro localizado no centro de um anel orgânico heterocíclico chamado de porfirina

Doenças genéticas ligadas ao metabolismo e transporte de lipídios

Hipercolesterolemia familial

A hipercolesterolemia familial apresenta herança autossômica dominante, diferentemente de todas as outras doenças anteriormente descritas, que eram de herança autossômica recessiva. Essa doença parece ter papel importantíssimo nos distúrbios causados por alterações fisiológicas ligadas ao metabolismo de gorduras ou lipídios. Calcula-se que 1 em 500 indivíduos são afetados por essa predisposição familiar, sendo responsável por cerca de 5% de todos os infartos do miocárdio. A hipercolesterolemia familial surge de mutações herdadas, localizadas no gene que codifica o receptor para lipoproteínas de baixa densidade ou LDL.

A Figura 7.3 apresenta um esquema da via do receptor de LDL e as várias etapas no ciclo de receptores.

O conjunto de moléculas de colesterol do organismo humano é formado pela absorção do colesterol da dieta e pela biossíntese nas células do fígado e intestino. Assim, as moléculas que são sintetizadas nesses órgãos devem ser transportadas para todas as células, como lipoproteínas. Os receptores para LDL são proteínas que estão presentes nas membranas celulares das células humanas e que têm o papel de ligar os complexos moleculares LDL presentes no sangue. As LDL ligadas ao receptor são engolidas pela célula por endocitose; depois, revestidas por uma proteína especial que sinaliza a formação de vesículas chamadas de endossomos, vão se juntar com lisossomos presentes nas células. Estes contêm diversas enzimas proteolíticas e vão digerir a parte proteica das lipoproteínas, liberando aminoácidos que poderão ser reutilizados para síntese de novas proteínas. Ao mesmo tempo, enzimas chamadas de colesterol esterases, também presentes nos lisossomos, vão liberar moléculas de colesterol para síntese de membranas, ácidos biliares, hormônios esteroides e novas lipoproteínas. As mutações presentes no gene do receptor de LDL são muito diversas e podem afetar sobremaneira a utilização das moléculas de colesterol. De modo geral, essas mutações vão afetar a captação das LDL pelas células, levando a um aumento dos níveis de colesterol sanguíneo, caracterizando a chamada hipercolesterolemia familial, predispondo, assim, os indivíduos ao infarto do miocárdio e doenças afins. O tratamento sempre envolve dieta de alimentos ricos em colesterol e terápicos que levem à diminuição dos níveis de colesterol circulantes.

Doenças genéticas ligadas ao metabolismo de porfirinas

As doenças conhecidas como porfirias estão associadas à produção alterada e herdada de um grupamento químico conhecido como heme, que contém um íon ferro e está presente nas proteínas ligadas ao processo respiratório, como as hemoglobinas, e ao processo de detoxificação, como os citocromos P450. Essas doenças genéticas levaram à descoberta da via de síntese do heme, mostrando que essa via é bastante complicada, já que as diferentes porfirias estão associadas às diferentes enzimas que fazem parte de sua biossíntese.

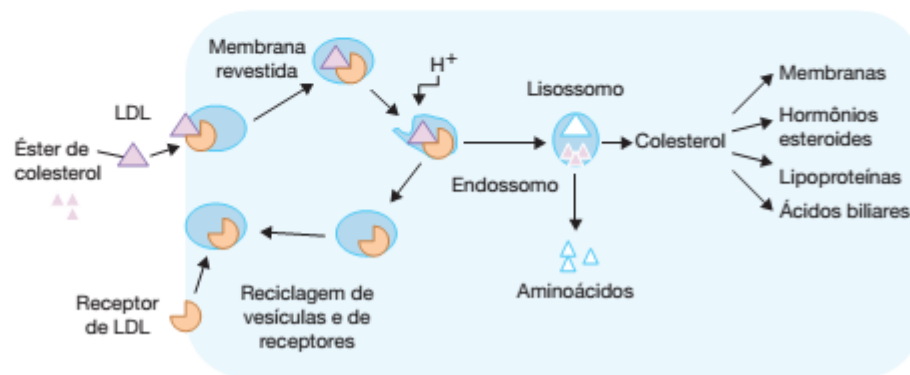



Figura 7.3 Via de metabolização da lipoproteína de baixa densidade – LDL por células humanas.

Porfirina
Classe de molécula orgânica formada por 4 anéis do tipo pirrol, ligados por pontes de carbono do tipo metínicas e cujo centro pode conter um íon metálico. Por causa das ligações duplas e intercaladas, essas moléculas absorvem luz visível, podendo, portanto, ser coloridas

Efeito do fundador
Diferença na frequência de determinado genótipo originado de um pequeno número de indivíduos portadores, ditos fundadores, em uma população isolada, diferindo da população de origem

As porfirinas são sintetizadas em todos os tecidos humanos, no entanto, o fígado e a medula óssea são os tecidos com o maior envolvimento em sua biossíntese. Essas observações seriam esperadas, já que esses tecidos estão intimamente ligados à produção de citocromos e hemoglobinas, respectivamente.

A Figura 7.4 apresenta um esquema da via de biossíntese do heme e as enzimas ligadas às porfirias e doenças associadas.

As porfirias, de modo geral, podem ser agrupadas em hepáticas ou eritropoéticas, dependendo se o acúmulo dos intermediários da síntese do heme ocorre predominantemente no fígado ou na medula óssea, respectivamente. Essas doenças têm como características clínicas gerais dores abdominais, transtornos psiquiátricos e fotossensibilidade cutânea. Podemos citar como importantes distúrbios do metabolismo de porfirinas, doenças genéticas autossômicas dominantes, como a porfiria *variegata*, a porfiria intermitente aguda e a porfiria cutânea tarda e, como autossômica recessiva, a porfiria eritropoética congênita. A porfiria *variegata* (OMIM 176200) é resultado da atividade reduzida da enzima protoporfirinogênio oxidase. Seu gene, o *PPOX*, está localizado no braço longo do cromossomo 1. Os indivíduos portadores dessa doença apresentam fotossensibilidade, alterações neurológicas e retardo no desenvolvimento, apesar de seu nome se dever aos diferentes quadros clínicos que podem surgir dessa condição. A porfiria *variegata* tem sido associada a um efeito do fundador na África do Sul. O Dr. Geoffrey Dean, que emigrou da Inglaterra para a África do Sul nos anos 1947, observou que algumas doenças tinham frequência maior nesse país. Ele se dedicou especialmente a essa porfiria e identificou um casal de holandeses colonizadores que lá se instalaram como sendo responsáveis pela distribuição da mutação. A porfiria *variegata* pode ser classificada como uma porfiria hepática. Já a porfiria eritropoética congênita (OMIM 263700), de caráter recessivo, é classificada como eritropoética, como o próprio nome diz. Ela é causada por mutações no

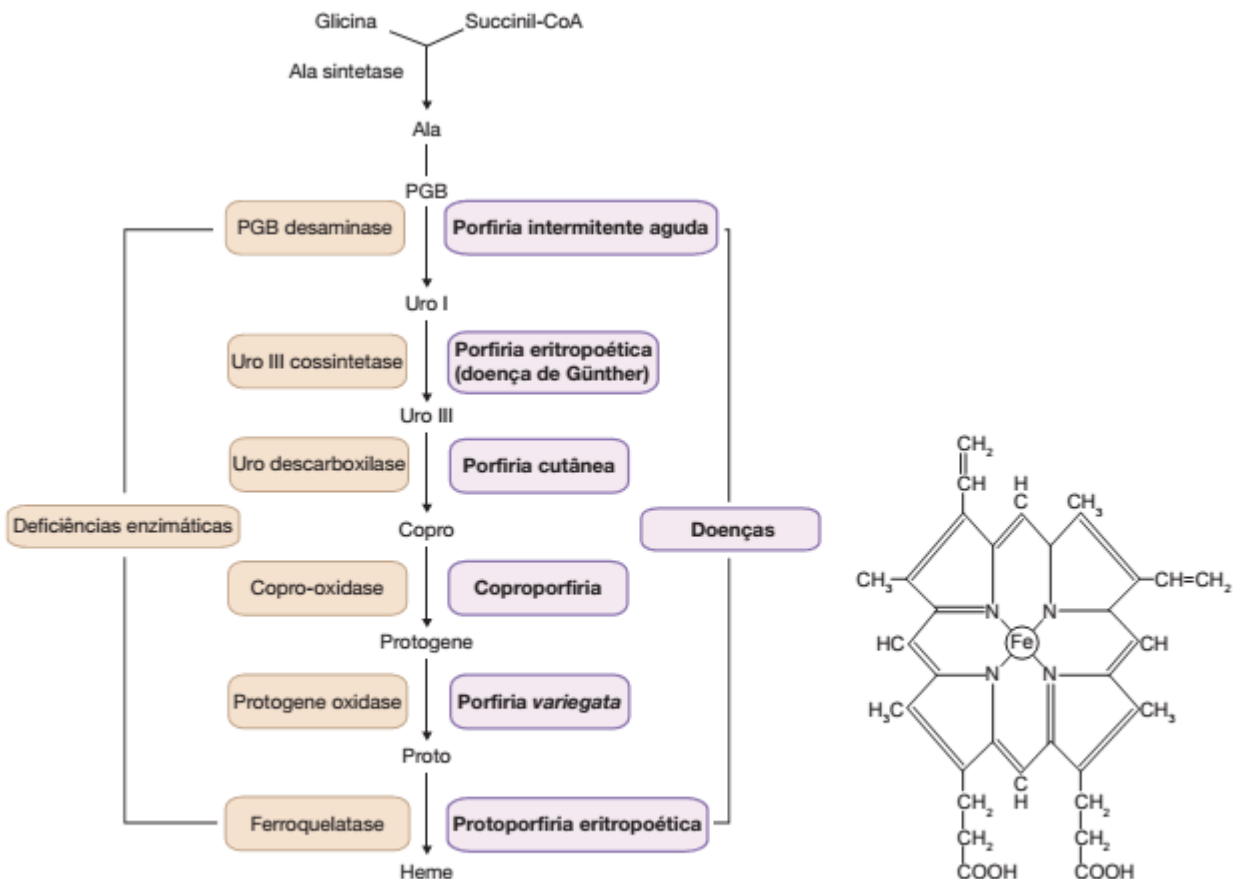



Figura 7.4 Via de biossíntese do heme e as enzimas ligadas às porfirias e doenças associadas.

gene da enzima uroporfirinogênio III sintase, localizado no braço longo do cromossomo 10. Diversas mutações foram descritas como responsáveis pela diminuição da atividade enzimática, o que resulta em efeitos bastante danosos. Essa doença é muito rara e os indivíduos afetados apresentam extrema fotossensibilidade e anemia hemolítica.

Hemoglobinopatias

Hemoglobina
Proteína presente nas hemácias, cujas funções principais são fazer o transporte de oxigênio dos pulmões para os tecidos e manutenção do pH interno constante

As hemoglobinopatias formam um conjunto de doenças que têm em comum alterações patológicas e herdadas na estrutura e funcionamento da proteína hemoglobina. A respiração é o processo pelo qual as células dependem da utilização de moléculas de oxigênio, gás abundante na nossa atmosfera, para o metabolismo e a produção de energia, produzindo o CO_2, que é liberado para o meio externo à célula. Nos organismos complexos como os vertebrados, esse sistema, para ser viável, depende da existência de moléculas para transporte do oxigênio. Portanto, a hemoglobina está no centro de todos esses processos em humanos, tendo participação fundamental na manutenção da nossa homeostase.

As doenças causadas pelas alterações genéticas resultando no funcionamento anormal da Hemoglobina podem, de modo geral, ser agrupadas em anemias e talassemias.

As anemias estão associadas a alterações por mutações presentes nos genes das cadeias polipeptídicas, produzindo proteínas anômalas, tanto estruturalmente quanto funcionalmente alteradas. Já as talassemias dizem respeito a alterações genéticas que levam à produção diminuída da hemoglobina funcional. Dessa pequena introdução, iremos descrever a estrutura, função das hemoglobinas e, finalmente, as doenças classificadas como hemoglobinopatias.

Estrutura e função da hemoglobina

Estrutura quaternária
Nível de organização de uma proteína cuja configuração é constituída pelo arranjo tridimensional de duas ou mais cadeias polipeptídicas

A hemoglobina é considerada uma daquelas moléculas cuja estrutura está totalmente associada ao seu desempenho funcional. Ela foi uma das primeiras proteínas a terem não só sua sequência de aminoácidos determinada, como também sua organização no espaço, ou seja, sua estrutura tridimensional. Não por acaso as descobertas a respeito da estrutura da hemoglobina foram tanto históricas quanto elucidativas para outras proteínas. A hemoglobina é um excelente exemplo de proteína cuja função depende de sua estrutura quaternária.

A hemoglobina no adulto é formada por 4 cadeias polipeptídicas, duas do tipo alfa e duas do tipo beta, cada uma associada a determinada estrutura conhecida como *heme*, que é um anel de porfirina contendo um átomo de ferro sob a forma iônica.

Esse íon ferro tem a capacidade de se associar temporariamente a uma molécula de oxigênio ou O_2. Portanto, cada molécula de hemoglobina pode carregar 4 moléculas de oxigênio, o que ela faz de maneira extremamente eficiente, transportando o O_2 dos pulmões aos tecidos que o utilizarão no metabolismo. Ainda, é importante ressaltar que o íon ferro tem de estar sob sua forma reduzida.

A Figura 7.5 mostra a representação da molécula da hemoglobina humana do adulto, com as cadeias constituintes e as estruturas do tipo heme nas quais o oxigênio molecular se liga.

Fato muito interessante sobre essa proteína é que as cadeias polipeptídicas que a formam podem existir de diferentes maneiras, produtos de diferentes genes que são expressos em diferentes ocasiões do desenvolvimento humano. Essas cadeias polipeptídicas variantes foram denominadas por letras gregas. Assim, existem hemoglobinas do tipo embrionária, formadas por duas cadeias do tipo alfa e duas do tipo épsilon – $\alpha_2\varepsilon_2$; fetal (Hb F), formada por duas cadeias do tipo alfa e duas do tipo gama – $\alpha_2\gamma_2$; e adulta (Hb A), formada por duas cadeias do tipo alfa e duas do tipo beta – $\alpha_2\beta_2$, como já mencionado. A Hb embrionária está presente durante a vida embrionária até 3 meses de vida intrauterina, após a qual a Hb F é majoritária até por volta do primeiro mês após o nascimento. A Hb A passa a estar presente por volta do mesmo período e passa a ser o tipo de hemoglobina funcionante durante toda a vida. Adultos normais também apresentam de 2% a 3% do tipo de Hb A2, que é composta por duas cadeias do tipo alfa e duas do tipo delta – $\alpha_2\delta_2$. Essas diferentes hemoglobinas têm constituição de aminoácidos um pouco diferente, o que reflete algumas propriedades fisiológicas diferentes. Por exemplo, a Hb F está adaptada ao ambiente onde o feto se desenvolve, ou seja, ao útero

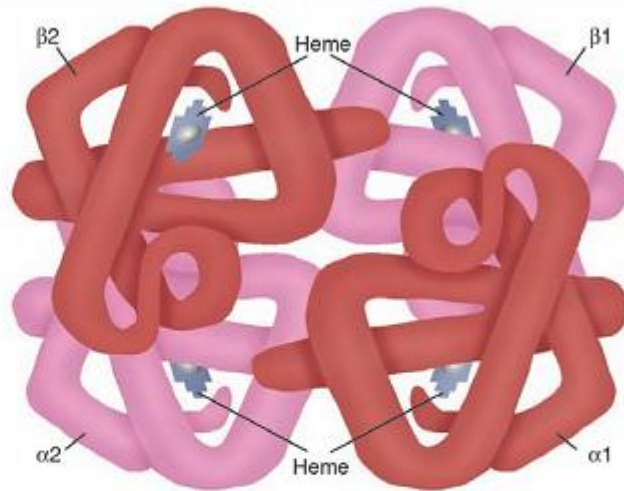



Figura 7.5 Representação da molécula da hemoglobina humana do adulto.

materno, onde o feto recebe o oxigênio do sangue da mãe que, certamente, é mais pobre nessa molécula do que a atmosfera. Portanto, a Hb F tem maior afinidade com o O_2, **podendo, então, competir com a Hb A materna e garantir o suprimento de O_2**. A afinidade da Hb pelo O_2 **pode** ser avaliada por curvas em um gráfico no qual a pressão do gás O_2 **é variável de 0 mmHg a 100 mmHg, e uma solução de Hb é avaliada quanto à sua capacidade de ligar o O_2, em termos de porcentagem de associação ou saturação. Lembre que cada molécula de Hb tem 4 locais de ligação, com o O_2** associando-se às 4 estruturas do tipo heme. Esse fato faz com que a curva seja **sigmoide, ou seja, tenha a forma da letra S, pois, depois da ligação com a primeira molécula de O_2**, a próxima se liga mais facilmente. A Figura 7.6 mostra o que acabamos de descrever a respeito da Hb F, mostrando a curva de saturação das hemoglobinas fetal e materna, do tipo sigmoide, e também, o deslocamento da curva da Hb F para a esquerda, ou seja, a Hb F torna-**se saturada de O_2 com menor pressão do gás.**

Assim, como o exemplo da Hb F, comentado anteriormente, as cadeias polipeptídicas, diferentes em sua sequência de aminoácidos, têm afinidades diferentes de associação ao O_2. **A seguir, discutiremos os genes globina e como eles são expressos, resultando na produção de diferentes polipeptídios.**

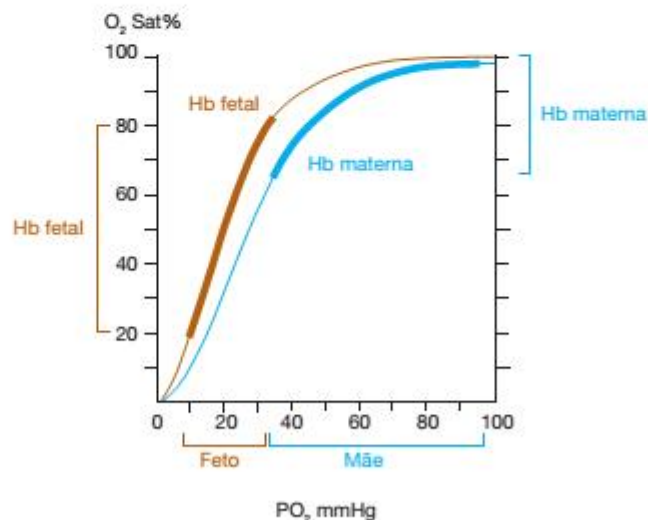



Figura 7.6 Curva de saturação das hemoglobinas fetal e materna.

Os genes globina, sua organização no genoma e a síntese da hemoglobina

Família gênica
Agrupamento de genes relacionados funcionalmente e que ocupam diversos *loci* no genoma. Acredita-se que os genes de uma família tenham surgido de um único gene ancestral, principalmente pela similaridade de suas sequências nucleotídicas

Pseudogenes
Genes similares a genes funcionais, sendo que perderam sua função em virtude de mutações

Os genes que produzirão as globinas, ou seja, as subunidades proteicas que vão formar a proteína tetramérica, podem ser de 2 tipos: um do tipo alfa e outro do tipo beta. Na verdade, formam famílias gênicas, sendo a família gênica do tipo alfa localizada no cromossomo 16 e a do tipo beta localizada cromossomo 11. Esses agrupamentos de genes são exemplos do que se conhece como famílias multigênicas de genes não idênticos mas relacionados estrutura e funcionalmente. Nas famílias dos genes globinas são encontrados os genes que serão expressos sequencialmente durante o desenvolvimento, assim como diversos pseudogenes.

Acredita-se que as famílias dos genes globina tenham sido resultado de forças evolutivas que, durante cerca de 450 milhões de anos, foram sofrendo mutações do tipo duplicação a partir de um gene ancestral comum. Assim, genes do tipo alfa e do tipo beta ancestrais sofreram uma série de duplicações e divergências em sua sequência de nucleotídios, tornando-se as duas famílias de genes que, como resultado desse processo evolutivo, mantiveram os pseudogenes como resquícios de eras passadas. Esses pseudogenes são transcritos mas não traduzidos, sendo, portanto, inativos. Ainda, a importante proteína mioglobina, cuja função principal é manter o suprimento de O_2 para o músculo esquelético, também faz parte dessa família de genes e, provavelmente, surgiu da mesma sequência de eventos evolutivos. A Figura 7.7 mostra um esquema que representa como os grupos gênicos possivelmente evoluíram de um gene ancestral comum.

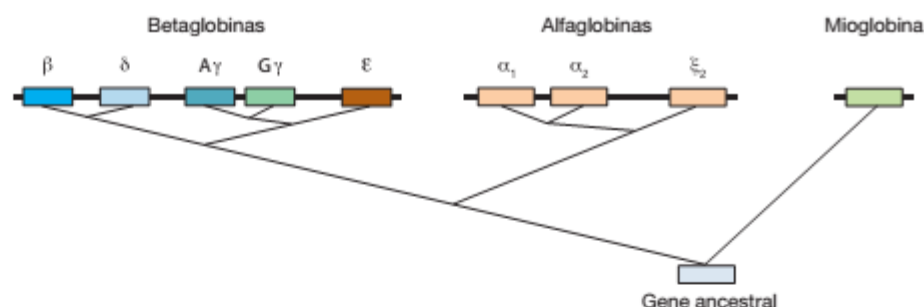



Figura 7.7 Possível evolução dos genes globina humanos.

Cada gene do grupo de genes globina, alfa ou beta, tem estrutura de gene ativo, ou seja, cada um com sua própria região promotora que, por sua vez, será controlada diferencialmente durante o desenvolvimento humano. Assim, como já foi mencionado, cadeias globina variantes vão aparecer no organismo em desenvolvimento nas diferentes fases. Mas como isso acontece? De modo bastante interessante, os genes de cada grupo alfa ou beta serão controlados por região controladora, localizada muitas quilobases acima do grupo gênico, chamadas de regiões antecedentes controladoras de *locus* ou LCRα, no grupo gênico alfa e LCRβ, no grupo beta. Essas regiões estarão associadas funcionalmente a sequências acentuadoras próximas das regiões promotoras de cada gene e respondem por fatores proteicos responsáveis pelas diferentes etapas do desenvolvimento humano, do embrião ao estado de recém-nascido. Na Figura 7.8, podemos observar um esquema da disposição do LCR dos genes de globina do tipo beta.

A Tabela 7.2 apresenta os diferentes tipos de hemoglobinas humanas, sua formação tetramérica com as cadeias globina variantes e o padrão de expressão durante o desenvolvimento humano.

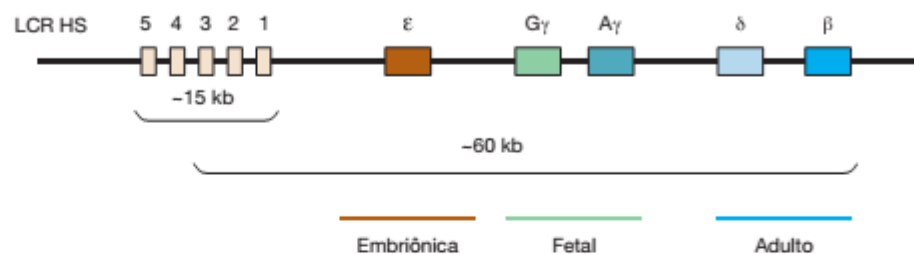



Figura 7.8 Localização das regiões antecedentes controladoras de *locus* ou LCR dos genes de globina do tipo beta.

■ **Tabela 7.2** As diferentes hemoglobinas e fase de desenvolvimento em que é expressa.

Hemoglobina	Estrutura molecular	Fase do desenvolvimento
Hb Gower I	$\zeta_2\varepsilon_2$	Embrionária
Hb Gower II	$\alpha_2\varepsilon_2$	Embrionária
Hb Portland	$\zeta_2\gamma_2$	Embrionária
Hb F	$\alpha_2\gamma_2$	Fetal
Hb A	$\alpha_2\beta_2$	Adulta
Hb A2	$\alpha_2\delta_2$	Adulta

Para a síntese da molécula completa, funcional da hemoglobina, são necessárias as várias etapas comuns a todas as proteínas sendo que, diferentemente de outras proteínas, as hemoglobinas têm de ter suas cadeias variantes expressas temporalmente, quantitativamente e localmente próximas, além de receber cada subunidade uma estrutura do tipo heme. Assim, é bastante complexa a sua via biossintética. As células precursoras dos eritrócitos, em geral denominados eritroblastos, presentes na medula óssea, constituem o principal sítio de biossíntese da hemoglobina. Após a síntese de toda a hemoglobina, essas células, em humanos, expulsam o núcleo e vão constituir as células vermelhas do sangue, chamadas eritrócitos ou hemácias. Essas células, na corrente sanguínea, são extremamente especializadas na função de transporte de gases, sendo praticamente como um pequeno saco de hemoglobina, circulante, cuja cor vermelha se deve ao heme da hemoglobina.

O metabolismo dos eritrócitos está voltado principalmente para a produção de energia, para sua sobrevivência e para manter o ferro do heme das hemoglobinas na forma reduzida ou íon ferroso, sem a qual o O_2 não poderá ser transportado. O tempo de vida médio das hemácias é de, aproximadamente, 120 dias.

Patologias associadas à hemoglobina alterada

As patologias associadas ao funcionamento anormal das hemoglobinas em humanos são chamadas de hemoglobinopatias. Muitas dessas anomalias são de origem hereditária e podem ser divididas em 2 grandes grupos: patologias causadas por alteração estrutural da Hb, na qual se encontram alguns tipos de anemias; e patologias causadas por alterações na síntese das cadeias globinas, ou talassemias. A Tabela 7.3 mostra as variantes genéticas presentes nas diferentes patologias e os tipos de mutações associadas.

Como observado na Tabela 7.3, existem muitas e diferentes variantes de hemoglobina que resultaram de mutações de tipos diferentes. Todas essas alterações levam a estados patológicos mais ou menos graves, de acordo com o tipo de mutação, ou seja, a mutação que causa uma alteração estrutural muito grande (ver Hb S, adiante) vai resultar em sintomas mais graves, mesmo estando em heterozigose. Outras alterações estruturais vão formar variantes cuja função da Hb não esteja tão prejudicada, causando sintomas mais brandos, como é o caso da Hb CS

Tabela 7.3 Patologias associadas à presença de hemoglobinas com estrutura alterada.

Patologia	Hemoglobina alterada	Descrição da alteração
Anemias falcêmicas	Hb S	Mutação de ponto, troca de ácido glutâmico por valina na cadeia beta (βGlu6Val)
	Hb C	Mutação de ponto, troca de ácido glutâmico por lisina na cadeia beta (βGlu6 lis)
Cianose por Hb de baixa afinidade pelo O_2	Hb M (Boston)	Mutação de ponto, troca de histidina por tirosina na cadeia alfa (αHis58Tir)
	Hb Kansas	Mutação de ponto, troca de asparagina por treonina na cadeia beta (βAsn102Thr)
Anemia por Hb instável	Hb Gun Hill	Deleção na cadeia betacódons 91 a 95
	Hb Koln	Mutação de ponto, troca de valina por metionina na cadeia beta (βVal98Met)
Policitemia por Hb com alta afinidade pelo O_2	Hb Heathrow	Mutação de ponto, troca de fenilalanina por leucina na cadeia beta (βPhe103 leu)
Talassemias	Hb Lepore	*Crossing over* desigual entre as cadeias δ e β, formando cadeia de fusão (δβ)
	Hb E	Mutação de ponto, troca de ácido glutâmico por lisina na cadeia beta (βGlu26 lis), ativação de sítio de recomposição
	Hb CS (Constant Spring)	Mutação de ponto, convertendo um códon de parada em códon para glutamina, resultando em cadeia alfa alongada (α142, Term->Gln)

e da Hb Lepore. Ainda pode haver mais de um tipo de mutação presente em um mesmo indivíduo, resultando em interações entre as hemoglobinas variantes. De modo interessante, esse tipo de ocorrência não é de todo prejudicial. Pode-se observar um exemplo nos indivíduos que herdaram mutações que produzem a Hb S e Hb C, ambas no códon 6 da cadeia β. Essa dupla mutação resulta em uma anemia falciforme mais branda.

Anemia falciforme

A anemia falciforme (OMIM 603903) é uma anemia hemolítica de caráter hereditário, apresentando herança autossômica recessiva. Talvez seja a anemia hereditária mais comum, com alta incidência em africanos e afrodescendentes, e tendo sido descrita pela primeira vez nos anos 1910 pelo médico James B. Herrik.

A mutação gênica responsável pela anemia falciforme é uma mutação de ponto, do tipo transversão. Aqui, ocorre a troca de uma adenina para uma timina, no sexto códon do gene da cadeia β, trocando o aminoácido ácido glutâmico por valina, formando a Hb S, com cadeias β alteradas.

Essa Hb tem baixa afinidade pelo O_2 e tende a se polimerizar nos homozigotos para a mutação. Além da inoperância do transporte de oxigênio, essa polimerização da Hb S tem consequências graves, induzindo eritrócitos a adquirirem a forma de foice. Esse afoiçamento das células vermelhas do sangue leva a uma série de distúrbios clínicos que resultam em morte precoce do indivíduo portador. São muitos os sintomas dessa patologia, como isquemia e trombose, por conta do aumento da viscosidade do sangue e de esplenomegalia e fraqueza, por conta da anemia. Os portadores heterozigotos para a mutação apresentam os chamados traços falcêmicos, sendo, em geral, saudáveis, podendo apresentar episódios da patologia durante a vida. Muito interessante é a proteção à malária em indivíduos portadores de heterozigose que têm, finalmente, menor suscetibilidade à doença. Isso resulta na persistência do alelo tipo S na África tropical, região onde a malária é muito comum.

O traço falcêmico se caracteriza pela presença de mutação para a anemia falciforme em heterozigose. Os portadores, em geral, não têm a doença, mas produzem parte da hemoglobina sob a forma HbS. Em 1954, A. C. Allison reportou que, segundo vários relatos, inclusive sua própria observação, a presença de alta frequência da mutação em populações africanas era

coincidente com a presença de malária. Por exemplo, Beet, em 1946, observou que em um grupo de 102 portadores de anemia falciforme de um distrito do norte da Rodésia, atual República de Zâmbia, apenas 9,8% apresentavam os parasitos da malária no sangue contra 15,3% do grupo de não portadores. Uma relação direta parece explicar essa observação, uma vez que as células contendo a HbS não são adequadas para a infestação pelo *P. falciparum* e sua reprodução. Assim, indíviduos portadores de traços falcêmicos são mais resistentes à malária e, portanto, sobrevivem em locais de alta frequência dessa doença, perpetuando a presença da mutação nas populações dessas regiões.

Talassemias

As talassemias são anemias resultantes da diminuição de globinas, tanto alfa quanto beta, por produção de globinas anômalas (ver, anteriormente, Hb Lepore e Hb CS) ou pela alteração na sequência da região controladora do genoma, região amplificadora por exemplo, ou ligada ao processamento do RNA mensageiro, nos sítios de recomposição nos genes globina. Esses eventos desequilibram a produção global de hemoglobina pela falta de uma das cadeias com o consequente distúrbio no transporte de oxigênio. A talassemia α (OMIM 141800) é causada principalmente por deleção de um ou mais genes de globina α, resultantes de *crossing over* desigual entre sequências homólogas na família de genes globina α. Isso pode ocorrer em virtude da grande similaridade entre os genes α_1 e α_2. Existem diferentes condições dessa patologia, dependendo do número de genes deletados no evento mutacional. Cada um de nós tem 4 cópias funcionais de genes globina α. Portanto, temos genótipo igual a 2 α/2 α. Nos indivíduos portadores de deleção de um gene (2 α/α -), os sintomas da talassemia não são percebidos, e diz-se que são portadores de traço falcêmico. Já na perda de 2 genes, genótipos α -/α - ou 2 α/- -, os portadores têm a apresentação clínica de uma anemia branda e, quando 3 genes são deletados (α -/- -), os sintomas são de anemia hemolítica moderada. Nessa condição, a Hb formada pode conter 4 monômeros de globina do tipo β (β_4), chamados de Hb H. Finalmente, com os portadores de deleção de todos os alelos, tem-se o tipo letal da talassemia α, vindo o portador a falecer antes de nascer. Essa condição patológica é chamada de hidropsia fetal, tendo o feto herdado a deleção genética obrigatoriamente dos 2 pais. Observou-se que a Hb formada por esses natimortos portadores era do tipo γ_4 ou $\zeta_2\delta_2$, sendo denominadas Hb Barts e Hb Portland, respectivamente. Já a talassemina β (OMIM 141900), de maneira interessante, resulta raramente de deleção gênica, mas, sim, de eventos mutacionais diversos que prejudicam a expressão das globinas do tipo β normal; por exemplo, mutações de ponto, deleções e inserções de algumas bases que podem estar localizadas em regiões codificadoras ou não codificadoras, mas controladoras. Por exemplo, mutações que afetam os sítios de encadeamento doadores ou receptores (Capítulo 2) dos íntrons dos genes globina β levam à recomposição anômala do transcrito e, portanto, a expressão do gene fica comprometida. Essa heterogeneidade de alterações genéticas resulta em talassemias do tipo β de diferentes gravidades, sendo classificadas em talassemia major ou β^0, intermediária e, ainda, talassemia minor. A talassemia major é a mais grave, também conhecida como talassemia de Cooley por ser esse o nome do médico; na verdade, um pediatra americano, o qual a descreveu nos anos 1920. Indivíduos homozigotos para a condição β^0 dependerão de transfusões sanguíneas regulares por toda a vida. O transplante de medula óssea parece ser uma opção de sucesso para o tratamento desses pacientes.

■ Imunogenética

A manutenção do meio interno constante em humanos também diz respeito aos mecanismos de defesa que têm como papel principal impedir a instalação de agentes infecciosos, causadores de doenças. Esses agentes podem ser vírus, bactérias, protozoários e fungos, e devem ser reconhecidos como estranhos e sofrer uma série de reações de neutralização da sua presença. Assim, é do equilíbrio entre o reconhecimento do que é próprio e não próprio, em uma espécie de vigilância fisiológica, que se estabelece o funcionamento adequado do organismo. O sistema

Interferons
Grupo de proteínas produzidas por células humanas teciduais ou linfócitos que têm a propriedade de induzir resistência viral às células não infectadas

Sistema complemento
Grupo de, aproximadamente, 30 proteínas que estão no sangue e que são ativadas por uma cascata de reações bioquímicas, produzidas por interações complexas

Células fagocitárias
Células que têm capacidade de fagocitar, ou seja, podem emitir prolongamentos celulares que vão englobar partículas sólidas e bactérias, promovendo sua destruição

Células citotóxicas naturais
Células conhecidas como *natural killers* ou NK, são linfócitos citotóxicos especializados em destruir células infectadas por vírus ou cancerosas

envolvido com essas reações biológicas é o sistema imunológico. A imunogenética é o ramo da genética que estuda e descreve as bases genéticas das reações imunológicas.

Tipos de defesas imunológicas

Nos animais vertebrados, incluindo o homem, os mecanismos imunológicos ocorrem em 3 fases essenciais: reconhecimento do elemento como não próprio; comunicação desse reconhecimento a células de resposta e eliminação do elemento estranho.

Existem 2 tipos de defesas imunológicas: uma dita *imunidade inata* e outra, conhecida como *imunidade adquirida*. A definição do termo inato significa "o que nasce conosco", portanto, esse tipo de imunidade envolve mecanismos que, supostamente, já estão presentes nos organismos, prontos para cumprir a função de defesa. Podemos enumerar como imunidade inata as barreiras de defesa como a pele, mucosas e secreções. Um bom exemplo é a existência da enzima lisozima na saliva. Esta tem como atividade digerir elementos glicídicos básicos presentes nas paredes bacterianas e, portanto, pode eliminar uma boa quantidade de bactérias presentes na boca. Outros 2 grupos de proteínas importantes nos sistemas de defesa inata são os interferons e o sistema complemento. Os interferons constituem a resposta primária à infecção viral, sendo produzidos pelas células infectadas, promovendo resistência viral às células não infectadas. Essa resistência é resultado do estímulo da produção pelas células não infectadas, de substâncias que inibam a reprodução dos vírus. Já o sistema complemento é formado pela presença de proteínas circulantes na corrente sanguínea, inativas, que serão ativadas pela presença de substâncias características da superfície de muitos microrganismos, levando à destruição destes. Apesar de ser importante na resposta imunitária inata, o sistema complemento também participa de reações inflamatórias e na resposta imunológica adquirida.

Também presentes como importantes defesas inatas do organismo estão as células com capacidade de fagocitar ou células fagocitárias (macrófagos, neutrófilos e células dendríticas), a resposta inflamatória e a presença no sangue de células especializadas conhecidas como *células assassinas naturais*, também conhecidas como células citotóxicas naturais ou *natural killers*. Por ser este capítulo dedicado aos mecanismos genéticos de adaptação e à especificidade da resposta imunológica, não entraremos em detalhe sobre o papel dos mecanismos de imunidade inata.

A resposta imunológica específica: o reconhecimento do próprio e do não próprio

A imunidade adquirida está associada a um tipo de imunidade também conhecida como adaptativa. Nesse tipo de imunidade é necessário que o organismo produza uma reação específica e dependente do tipo de agente que é reconhecido como não próprio do organismo. A base para esse tipo de imunidade é justamente a distinção correta entre o que é próprio e o que não é próprio. Assim, o recrutamento de elementos celulares ou moleculares é realizado para a eliminação de microrganismos, de células estranhas e até mesmo de células cancerosas.

Receptores celulares
Moléculas de natureza essencialmente proteica que interagem com substâncias externas às células, desencadeando alterações nestas

Antígenos
Estruturas capazes de induzir a resposta imunológica específica

A principal característica da imunidade adquirida é sua especificidade. As principais células envolvidas são os linfócitos B e T, e as principais proteínas envolvidas são os receptores celulares, presentes nas células B e T, e as imunoglobulinas. A resposta imunológica adquirida vai, então, ser dirigida contra determinada estrutura reconhecida como não própria, chamada de antígenos.

Os linfócitos B e T são células presentes na circulação sanguínea produzidas na medula óssea a partir de células-tronco. Essas células se tornarão células maduras por meio da *diferenciação*. Os linfócitos T vão se diferenciar em um local específico do corpo humano, um órgão chamado timo, e os linfócitos B vão se diferenciar na própria medula óssea. Portanto, a designação "T" vem de timo e "B" vem de bursa de Fabricius, um órgão que foi descoberto nas aves com função semelhante à da medula óssea nos mamíferos. A imunidade adquirida parece ser exclusividade dos vertebrados. Na verdade, o sistema imunológico dos vertebrados é altamente complexo e interativo e as respostas imunológicas do tipo inata e adquirida são, muitas vezes, interdependentes.

Citocinas
Moléculas existentes no meio extracelular que atuam na transmissão de sinais entre as células

Complexo principal de histocompatibilidade ou MHC
Família de genes que codificam diferentes proteínas de superfície celular, fornecendo identidade às células de um mesmo organismo e fazendo parte do processo de apresentação de antígenos

Linfócitos T auxiliares
Células T que respondem à apresentação do antígeno pelas células apresentadoras de antígeno ativando linfócitos B a produzir anticorpos e linfócitos T a produzir receptores específicos, induzindo tanto a resposta imunológica humoral quanto a celular

Linfócitos B
Células que se originam na medula óssea e vão se diferenciar nesse mesmo local, mas, segundo estímulos, podem formar os plasmócitos, que são as células circulantes produtoras de anticorpos

Linfócitos T
Células que se originam na medula óssea e vão se diferenciar no timo, e, segundo os estímulos, vão formar diferentes tipos de linfócitos T

Quando um organismo entra em contato com elementos estranhos, os macrófagos e as células dendríticas vão ingeri-los e processá-los.

Após a etapa de processamento, essas células irão expor fragmentos das estruturas ingeridas com as moléculas que existem na sua superfície, formando complexos proteicos. Esse conjunto de eventos é chamado de apresentação de antígenos e as células que têm essa capacidade são chamadas de células apresentadoras de antígenos.

Células especializadas serão recrutadas pela liberação de citocinas e, assim, a resposta imunológica se desenvolve.

A Figura 7.9 apresenta os eventos de ativação das células imunológicas na resposta, por exemplo, a uma infecção por bactéria.

As moléculas que estão localizadas na superfície das células que participam da formação dos complexos de apresentação de antígenos são conhecidas como proteínas do complexo principal de histocompatibilidade ou MHC.

O complexo principal de histocompatibilidade ou MHC é um conjunto de genes que codificam proteínas de superfície. Essas proteínas, além de formarem um complexo com o antígeno processado, também são encarregadas de fornecer identidade às células de um mesmo indivíduo, que, então, são reconhecidas como próprias do organismo.

Uma vez formado esse complexo de proteínas do MHC e antígeno, os macrófagos vão ativar linfócitos T auxiliares ou célula T *helper*, por meio de interações com receptores. Para promover a resposta imunológica, os receptores devem, então, se ligar simultaneamente à proteína do complexo de histocompatibilidade (próprio) e ao antígeno (não próprio). Essas células T auxiliares vão ativar linfócitos B e T que, uma vez ativados, irão reconhecer e dirigir a resposta para os antígenos de maneira específica.

Assim, o MHC tem papel central na resposta imunológica. As proteínas do MHC que são denominadas em humanos de antígenos HLA, do inglês *Human Leucocyte Associated*, foram inicialmente identificadas nas respostas de rejeição a transplantes. No entanto, ao se estudar mais detalhadamente o papel destas proteínas, foi observado que essas moléculas estão intimamente associadas ao reconhecimento do que é próprio e não próprio ao organismo em um sentido mais amplo. O *locus* genômico do MHC, ou *locus* HLA, em humanos, corresponde a uma grande região genômica, localizada no cromossomo 6, que tem genes que codificam para 3 classes principais de proteínas do MHC. Em geral, as proteínas HLA são proteínas de membrana que apresentam uma porção extracelular, exposta, uma porção transmembranar e uma intracelular. Os genes do *locus* HLA são muito polimórficos, contendo, em alguns casos mais de 100 alelos diferentes, o que contribui para grande diversidade de moléculas entre indivíduos. Assim, os genes da classe I de HLA codificam para proteínas chamadas de antígenos de transplantes por estarem ligados à rejeição de tecidos nos transplantes. Os genes de classe II codificam proteínas presentes na superfície de linfócitos B e macrófagos, participando, portanto, da interação com os linfócitos auxiliares na resposta imunológica. Já os genes de classe III codificam as proteínas do sistema complemento, anteriormente citado como importante processo da imunidade inata

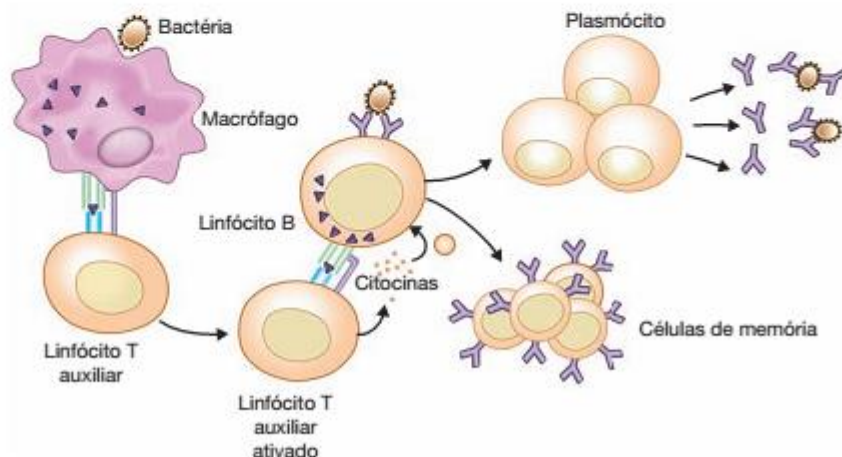



Figura 7.9 Eventos de ativação das células imunológicas na resposta a uma infecção por bactéria.

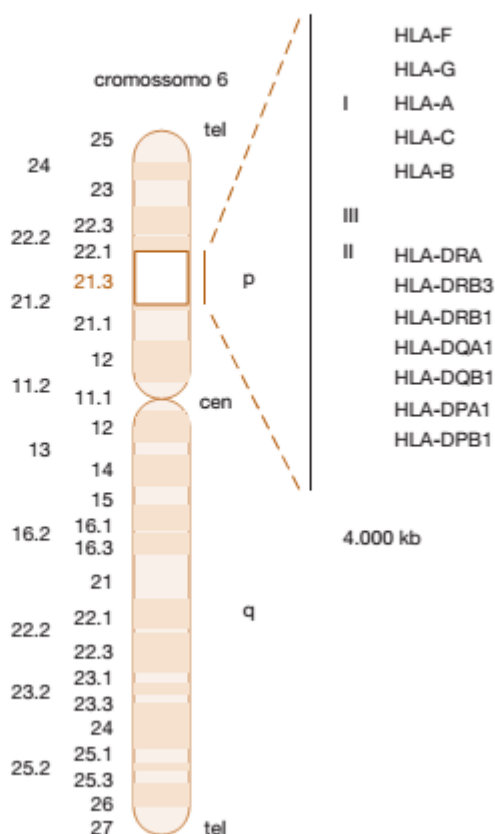



Figura 7.10 *Locus* HLA e sua localização no cromossomo 6.

e também participante da resposta imune adquirida. A Figura 7.10 mostra o cromossomo 6 e a representação de alguns genes HLA pertencentes às 3 classes e sua posição.

As células apresentadoras de antígenos realizarão esse papel por meio da interação do antígeno com as proteínas MHC do tipo II. Como já foi citado, essa interação resulta na ativação dos linfócitos auxiliares que estimulam os linfócitos T a expressarem receptores específicos e linfócitos B a se transformarem em plasmócitos, produtores de anticorpos, resultando nas respostas imunológicas do tipo celular e humoral, respectivamente. Tanto os receptores dos linfócitos quanto os anticorpos produzidos contra os antígenos têm enorme diversidade, podendo ser dirigidos para quase todo e qualquer antígeno que um organismo venha a estar exposto.

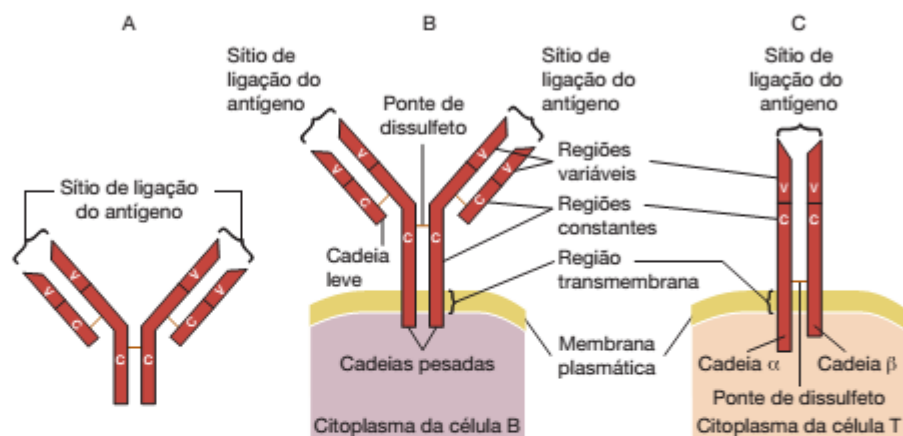
As moléculas de reconhecimento: os receptores de células T e B e as imunoglobulinas

Imunoglobulinas
Proteínas produzidas pelos linfócitos B diferenciados em plasmócitos, que atuam ao se ligar a determinantes antigênicos e que ajudam em sua eliminação; são também denominados anticorpos

De modo bastante interessante, os receptores das células B têm estrutura igual às **imunoglobulinas** e são considerados o tipo solúvel destes.

> As imunoglobulinas e os receptores das células B consistem em proteínas em forma de "Y", constituído por duas cadeias pesadas e duas cadeias leves, similares, ligadas por pontes dissulfeto. Os receptores de células T consistem em duas cadeias, uma do tipo α e uma do tipo β, ligadas por pontes dissulfeto.

Esses 3 tipos de moléculas têm em comum em suas estruturas uma porção constante, localizada na região C-terminal, e porções variáveis, localizadas na região N-terminal. Essas últimas são responsáveis pela enorme diversidade de moléculas e, portanto, pela especificidade da resposta imunológica. A porção variável está sempre localizada na região na qual tanto os receptores quanto as imunoglobulinas vão se ligar aos antígenos. A Figura 7.11 apresenta, esquematicamente, as estruturas dos receptores e da imunoglobulina.

Figura 7.11 Representação esquemática das estruturas de imunoglobulina (A) e de receptores de célula B (B) e célula T (C).

Toda diversidade encontrada nos receptores e nas imunoglobulinas acontece por processos genéticos de rearranjo genômico (ver capítulos anteriores). A seguir, iremos discutir a formação das imunoglobulinas para entendermos como se dá essa diversidade molecular. Apesar de apresentarmos apenas a construção de uma cadeia leve, temos de ter em mente que tanto os outros genes das cadeias das imunoglobulinas nas células B, como os receptores das células B e T sofrem processos muito semelhantes. As cadeias leves das imunoglobulinas são de 2 tipos diferentes chamados de cadeias kapa (κ) e cadeias lambda (λ), sendo o tipo determinado pela porção constante. Já as cadeias pesadas serão de 5 tipos diferentes e promovem a classificação das imunoglobulinas. Assim, a imunoglobulina do tipo A tem cadeia pesada do tipo α, e é chamada IgA. Assim, IgD com cadeia pesada δ; IgE com cadeia pesada ε; IgG com cadeia pesada γ e IgM com cadeia pesada μ. A Tabela 7.4 mostra as imunoglobulinas humanas em que estão presentes e sua função na resposta imunológica.

As imunoglobulinas vão formar os anticorpos que são considerados a resposta imunológica solúvel ou humoral, ou seja, são produzidos pelos plasmócitos e liberados no meio no qual irão atuar. Em relação à estrutura das Igs, além das cadeias leves e pesadas, podemos distinguir regiões funcionais ou domínios, presentes em todos os tipos de Igs, como os domínios de ligação ao antígeno, formados pelas regiões variáveis das duas cadeias leves e pesadas, e o domínio funcional efetor que é formado pelas regiões constantes das cadeias pesadas. Esse último é responsável pela interação com os outros elementos participantes da resposta imunológica.

Um anticorpo funcional vai se ligar a antígenos por meio de seus 2 domínios de ligação de modo tão específico que é comum chamar essa interação de tipo "chave-fechadura".

■ **Tabela 7.4** As diferentes classes de imunoglobulinas humanas.

Classe de Ig	Tipos de cadeia	Localização e características	Função
Ig A	α	Secreções como leite materno, saliva, lágrima e mucosas	Protege contra microrganismos em locais expostos ou de entrada
Ig D	δ	Principalmente em células B que não foram expostas a antígenos	Funciona como receptor de células B; a função não é definida
Ig E	ε	Normalmente no sangue em pequenas quantidades e em mastócitos	Participa da resposta da resposta inflamatória induzindo a liberação de histamina pelos mastócitos
Ig G	γ	Ig mais abundante no sangue, constituindo 80% das Igs	É a principal Ig da resposta imunológica para neutralização e eliminação de antígenos; participa da transferência de imunidade da mãe para o feto
Ig M	μ	Células B	Participa da fase inicial da resposta imunológica; muito eficiente em ativar o sistema complemento

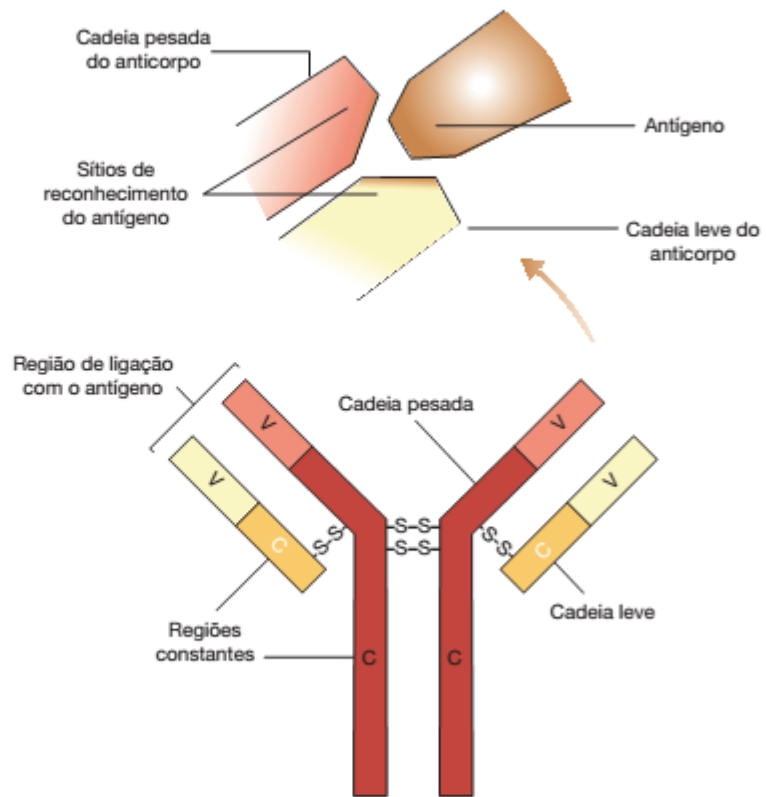



Figura 7.12 Esquema da estrutura da imunoglobulina apresentando o detalhe da ligação antígeno-anticorpo.

A estrutura específica do antígeno que vai promover a formação do anticorpo e que vai se ligar ao antígeno chama-se epítopo ou determinante antigênico.

A Figura 7.12 apresenta, esquematicamente, a estrutura da imunoglobulina e, em detalhe, a ligação antígeno-anticorpo.

Os anticorpos, como toda proteína, são produzidos por meio da expressão gênica que, no caso dessas proteínas tão especiais, acontece de maneira segmentada. A informação codificadora para a formação das cadeias polipeptídicas das Igs estará presente em diferentes *loci*, inclusive em cromossomos diferentes.

Os segmentos gênicos vão, então, ser alinhados, conforme o estímulo, em sequências adequadas por rearranjos genômicos, por eventos de recombinação que ocorrerão durante a diferenciação dos linfócitos B.

As células-tronco, presentes na medula óssea, têm em seu material genético toda a informação para a construção de cadeias das Igs. Para a construção de uma cadeia leve, por exemplo, essas células primordiais têm 3 tipos de segmentos: segmentos variáveis (V), segmentos de junção (J) e segmento constante (C). Os segmentos V e J, quando juntos, formarão a porção variável da cadeia leve. Já o segmento C codifica toda a porção constante. Por meio de técnicas de sequenciamento de bases do genoma humano, observou-se que existem 40 genes diferentes que codificam segmentos V, 5 genes para J e um gene para o segmento C. Assim, os segmentos podem ser combinados para formar cadeias leves de 200 formas diferentes para um mesmo tipo, κ ou λ. Resumidamente: um grupo de enzimas chamadas de recombinases irá catalisar a ligação entre um segmento V e um segmento J. Toda a região de DNA genômico entre esses 2 segmentos é, assim, eliminada. Eventos de encadeamento ou *splicing* serão responsáveis pela retirada da região intrônica existente entre esses segmentos colados e a porção C. O RNA maduro segue, então, para ser traduzido na cadeia polipeptídica de uma cadeia leve formada por uma porção variável e outra constante. A Figura 7.13 mostra o exemplo da montagem da cadeia polipeptídica da cadeia leve funcional.

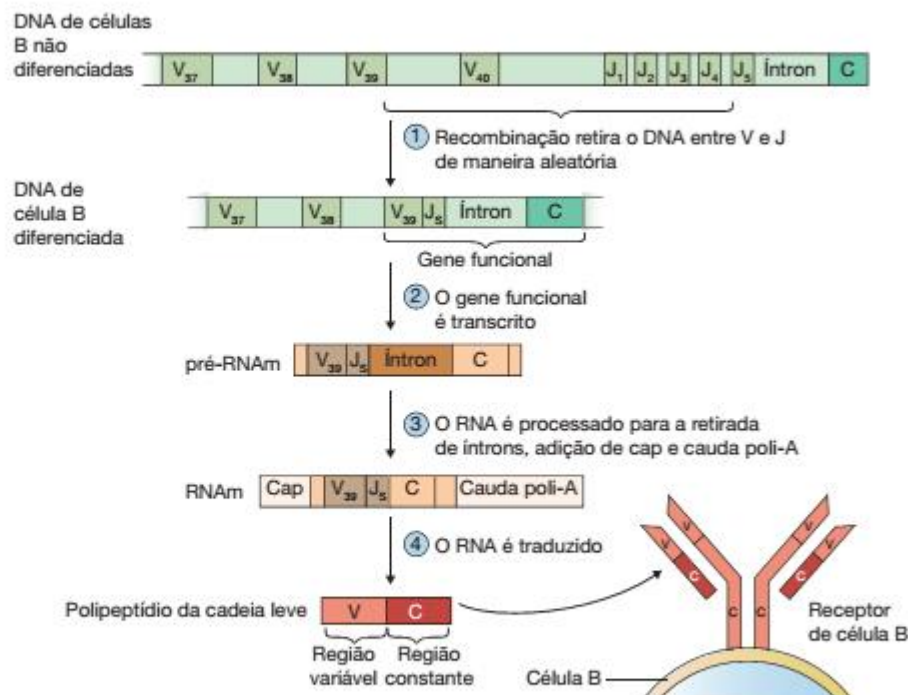



Figura 7.13 Montagem da cadeia polipeptídica da cadeia leve funcional.

Essa explanação do que acontece é bastante simplificada, pois tanto genes quanto recombinações gênicas durante o processo de diferenciação da célula B ou T são processos muito complexos. Por exemplo, as cadeias leves κ e λ, que estão localizadas em cromossomos diferentes, ou seja, cromossomo 2 e 22, respectivamente, têm número de segmentos diferentes entre si. Já a cadeia pesada tem segmentos do tipo D, de diversidade, além dos segmentos V, J e C. Lembre que as cadeias pesadas podem ser de 5 tipos diferentes! Não entraremos em detalhes, mas existem mais 2 eventos mutacionais que colaboram para a diversidade observada. O primeiro é chamado de diversidade juncional, por meio de sítios alternativos de recombinação durante as junções e o outro chamado de hipermutação somática. Assim, pode-se dizer que a natureza se reorganiza para que essas células especializadas tenham a capacidade de produzir moléculas direcionadas a um epítopo, produzindo a vigilância imunológica. Pode-se dizer que ao observarmos duas células B ou T maduras, ao acaso, dificilmente estas serão iguais, apresentando os mesmos receptores.

A resposta imunológica adquirida: uma visão geral

Depois de compreendermos as principais células e moléculas envolvidas na resposta imunológica adquirida, assim como os mecanismos que originam a especificidade e o reconhecimento do próprio e não próprio, podemos ter uma visão mais global da resposta imunológica. Quando o organismo humano entra em contato com um elemento exógeno pela primeira vez, este é fagocitado e processado por células apresentadoras de antígeno que, pela interação com as proteínas do MHC, vão ativar células T auxiliares. Essas células especializadas estimularão as células B e T a se diferenciarem e formarem células comprometidas com o reconhecimento, ataque e destruição desses elementos, agora considerados não próprios do organismo. Assim, as células B vão formar os plasmócitos, circulantes, que produzirão anticorpos específicos, resultando na resposta imunológica humoral. Já as células T vão formar as células T citotóxicas, originando a resposta imunológica celular.

Os eventos que resultam na formação dos plasmócitos e linfócitos citotóxicos produzem a primeira resposta do organismo chamada de resposta imunológica primária.

Células auxiliares de memória
Subtipos de linfócitos B ou T produzidos após a resposta imunológica primária

De modo bastante interessante, esses linfócitos T auxiliares formarão as células auxiliares de memória, ou seja, uma vez em contato com o elemento exógeno pela primeira vez, estes já

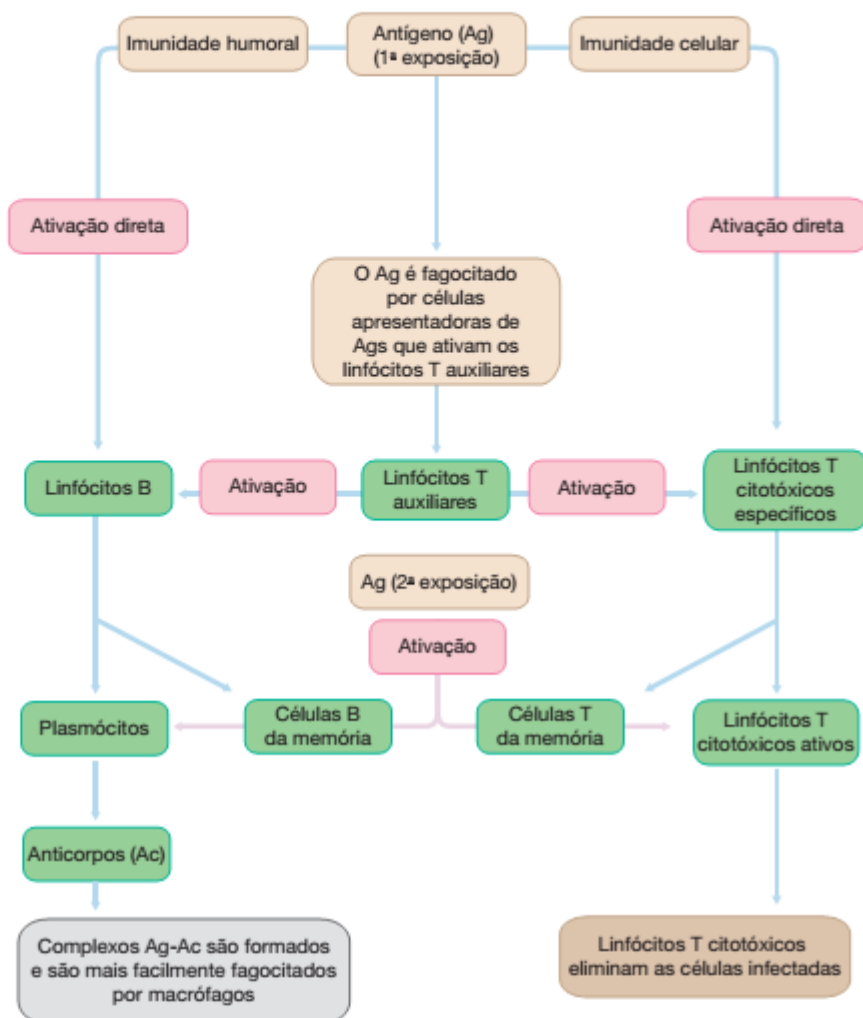



Figura 7.14 Esquema representando todas as etapas necessárias da resposta imunológica, da resposta primária e secundária.

vão produzir uma resposta mais imediata quando ocorrer uma segunda exposição. As células auxiliares de memória vão estimular a produção de grupos de células de memória do tipo B e células de memória do tipo T, prontas para a resposta imunológica. Esse tipo de resposta mais rápida do organismo é chamado de resposta imunológica secundária. A Figura 7.14 mostra todas essas etapas de maneira resumida.

Vacina
Substância obtida de um agente infeccioso que estimula a resposta imunológica direcionada a este, não produzindo a doença

Imunização ativa
Tipo de resposta imunológica que diz respeito à produção de anticorpos ou linfócitos T pelo organismo quando este entra em contato com o agente infeccioso

Imunização passiva
Tipo de resposta imunológica que diz respeito à produção de anticorpos ou linfócitos T por um organismo e transferida para outro

Aplicações e patologias relacionadas com a resposta imunológica adquirida

A defesa imunológica é considerada um processo fundamental para a manutenção do organismo funcional e, portanto, o meio interno equilibrado com o meio externo. Por meio de processos de imunização, mais comumente chamados de vacinas, milhares de indivíduos podem se proteger contra agentes patogênicos. Todos os mecanismos discutidos no item anterior fazem parte da imunização ativa. Já a imunização passiva se faz pela passagem de anticorpos, principalmente do tipo IgG, mas também do tipo IgA, de mãe para filho, pela placenta, no primeiro caso, ou pelo leite materno, no último caso. A vacinação está relacionada com a imunização ativa. A primeira vacina surgiu da utilização do vírus da varíola bovina, que promovia a formação de imunidade para o vírus muito mais patogênico para humanos, o vírus da varíola humana. Hoje, diversas fontes de antígenos são utilizadas na produção das vacinas, tais como toxinas

inativadas de microrganismos, fragmentos destes e ainda, a produção de proteínas antigênicas por técnicas de biologia molecular. Todos esses agentes promovem resposta imunológica primária e vão assegurar a produção de células T e B de memória, garantindo a imunidade pela resposta secundária.

A produção de anticorpos específicos também é muito utilizada em pesquisa para identificação e análise de proteínas importantes. O desenvolvimento de diagnósticos precisos para uma série de doenças, entre elas a AIDS (síndrome da imunodeficiência adquirida), se faz pela produção de *kits* de identificação. Embora tudo o que foi discutido mostre as propriedades benignas do sistema imunológico, muitas patologias surgem de distúrbios nesse sistema. As alergias, por exemplo, mostram que, no caso da resposta imunológica, o equilíbrio da produção de anticorpos é importantíssimo. As reações alérgicas, que são respostas exageradas da imunidade humoral contra antígenos, denominados alergênios, podem provocar reações inflamatórias sérias, podendo causar a morte pelo que se chama choque anafilático, que é uma resposta alérgica aguda. Na Figura 7.15 podemos observar o esquema representativo da interação entre a IgE produzida pela resposta imunológica e um mastócito, produtor de histamina.

Alergia
Resposta imunológica exacerbada dirigida a substâncias estranhas ao organismo

Choque anafilático
Resposta imunológica exagerada e aguda no indivíduo quando este entra em contato pela segunda vez com a substância estranha que promove reações patológicas importantes como inconsciência, convulsões e edema de glote

Doenças autoimunes
Grupo de doenças provocadas por disfunção do sistema imunológico, que passa a produzir resposta imune contra elementos próprios do organismo

Outro desequilíbrio importante da resposta imunológica diz respeito ao desenvolvimento das chamadas doenças autoimunes que, como o nome mostra, são causadas pela alteração do reconhecimento do que é próprio e não próprio, promovendo em alguns indivíduos uma resposta imunológica direcionada para elementos do organismo. A doença conhecida como lúpus eritromatoso sistêmico é uma doença autoimune com efeitos devastadores. O organismo de certos indivíduos passa a produzir anticorpos contra as histonas e DNA, expostos em virtude de quebras das moléculas pelos processos fisiológicos normais de destruição de células. A artrite reumatoide, a diabetes melito tipo I e a esclerose múltipla são exemplos de doenças autoimunes que afetam os seres humanos. A razão para que certos indivíduos passem a desenvolver autoimunidade está associada ao processo de autotolerância que faz com que linfócitos imaturos, produzidos durante a diferenciação, sejam eliminados por apoptose ou perda de função. Aparentemente, esses linfócitos têm receptores contra moléculas próprias do organismo, podendo causar danos a células e tecidos. Apesar do pouco conhecimento sobre essa etapa da resposta imunológica, sabe-se que fatores ligados à genética do indivíduo e ao ambiente são muito importantes. Ainda, como exemplo de resposta indesejada dos mecanismos imunológicos estão as reações de rejeição imunológica, observadas em transplantes de órgãos e tecidos e também nas transfusões sanguíneas. O evento da rejeição acontece porque células de outros indivíduos são consideradas elementos patogênicos estranhos, como se fossem bactérias e vírus. Implantes de tecidos de indivíduos não idênticos permanecerão inalterados por aproximadamente 1 semana quando, então, começa a destruição em massa das células, provocando reações inflamatórias de rejeição. Como já foi citado, as reações de rejeição a implantes de tecidos e transplantes de órgãos passam pela grande diversidade das proteínas do MHC. Para diminuir esse tipo de reação, a medicina atualmente procura indivíduos doadores que tenham a maior compatibilidade possível de moléculas MHC com o receptor.

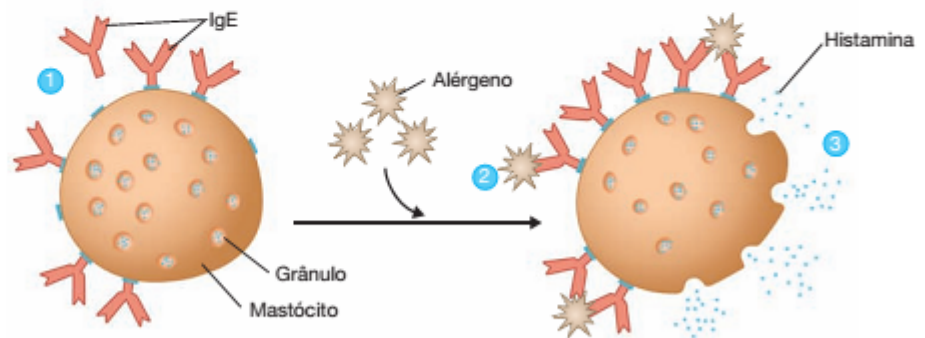



Figura 7.15 Interação entre a IgE produzida pela resposta imunológica e o mastócito.

RESUMO

- Doenças genéticas do metabolismo, do transporte de oxigênio e da resposta imunológica estão entre as mais estudadas e correspondem a um grande número de descobertas da era genômica
- As doenças genéticas do metabolismo são também conhecidas pelo termo erros inatos do metabolismo, que são doenças de caráter hereditário, causadas pela deficiência de proteínas pertencentes a vias metabólicas, resultante de mutações presentes em seus genes
- As doenças genéticas do metabolismo são agrupadas segundo a via metabólica a qual pertence e o tipo de herança; por exemplo, a hipercolesterolemia familiar é doença genética da via de metabolismo dos lipídios e é de herança autossômica dominante. Já a fenilcetonúria é ligada à via de metabolismo de aminoácidos e é de herança autossômica recessiva
- O banco de dados de doenças genéticas, denominado OMIM®, é uma importante fonte de informação a respeito das doenças genéticas de herança mendeliana
- Das doenças do metabolismo de aminoácidos destacam-se a alcaptonúria e a fenilcetonúria, sendo a alcaptonúria a primeira doença a ser estudada e classificada como erro inato do metabolismo
- A fenilcetonúria clássica, ou PKU, é uma doença genética causada por erros inatos na via de degradação do aminoácido fenilalanina. Ela produz uma hiperfenilalaninemia, que é o excesso de fenilalanina no organismo e que resulta em retardo mental grave nas crianças não tratadas, eczemas e cor clara de pele e olhos provocada pela falta da pigmentação. Essa doença teve importante papel no avanço das análises bioquímicas para recém-natos, motivando o desenvolvimento do teste do pezinho
- Outras doenças genéticas do metabolismo importantes são associadas ao metabolismo de carboidratos, como a galactosemia e as doenças de armazenamento de glicogênio, ao metabolismo de purinas e pirimidinas, como imunodeficiência congênita severa causada por deficiência de ADA, às doenças de metabolismo de lipídios, como a hipercolesterolemia familiar, e às doenças do metabolismo de porfirinas, como as porfirias
- As porfirinas envolvem a formação do radical heme, que participa da estrutura de inúmeras proteínas, entre elas a hemoglobina, especializada no transporte de oxigênio para os tecidos
- As hemoglobinopatias formam um conjunto de doenças que têm em comum alterações patológicas e herdadas na estrutura e funcionamento da hemoglobina
- A hemoglobina no adulto é formada por 4 cadeias polipeptídicas, duas do tipo alfa e duas do tipo beta, cada qual associada a uma estrutura do tipo heme, cujo íon ferro tem a capacidade de se associar temporariamente a uma molécula de oxigênio. Portanto, cada molécula de hemoglobina pode carrear 4 moléculas de oxigênio dos pulmões aos tecidos
- Os genes globina do tipo alfa e do tipo beta constituem famílias multigênicas, localizadas nos cromossomos 16 e 11, respectivamente. Esses agrupamentos gênicos têm genes que são expressos sequencialmente durante o desenvolvimento, controlados por regiões controladoras denominadas LCR, assim como diversos pseudogenes. Portanto, serão produzidas hemoglobinas diferentes durante a fase embrionária, fetal e adulta
- As doenças causadas pelas alterações genéticas, que resultam no funcionamento anormal da hemoglobina, podem, de modo geral, ser agrupadas em anemias e talassemias. As anemias estão associadas a alterações por mutações presentes nos genes das cadeias polipeptídicas, produzindo proteínas anômalas, funcionalmente alteradas. Já as talassemias dizem respeito a alterações genéticas que levam à produção diminuída da hemoglobina funcional
- A anemia falciforme, assim denominada em virtude da forma de foice apresentada pelas hemácias de indivíduos afetados, é uma anemia hemolítica de caráter hereditário, com herança autossômica recessiva e alta incidência em africanos e afrodescendentes. Ela é causada por uma mutação de ponto, com a troca do aminoácido ácido glutâmico por valina, formando a Hb S
- As talassemias podem ser do tipo alfa ou beta. A talassemia do tipo alfa é causada principalmente por deleção de um ou mais genes de globina alfa, sendo a hidropsia fetal um tipo grave de talassemia alfa, na qual os portadores têm deleção de todos os alelos da globina alfa e não sobrevivem. A talassemia tipo beta ocorre por eventos mutacionais diversos, como mutações de ponto, deleções e inserções localizadas em regiões codificadoras ou regiões controladoras. A talassemia major é um tipo grave de talassemia beta, sendo conhecida como talassemia de Cooley
- A imunogenética é o ramo da genética que descreve as bases genéticas das reações imunológicas em que agentes infecciosos (vírus, bactérias, protozoários e fungos) são reconhecidos como estranhos e sofrem uma série de reações para sua eliminação
- Existem 2 tipos diferentes de defesas imunológicas: a imunidade inata e a imunidade adquirida, sendo esta última específica para cada elemento considerado estranho
- Na imunidade adquirida, as principais células envolvidas são os linfócitos B e T, e as principais proteínas envolvidas são os receptores celulares presentes nas células B e T e as imunoglobulinas
- Os antígenos são estruturas capazes de induzir a resposta imunológica específica e formam os principais elementos estimuladores da resposta imunológica. Os antígenos ativam a resposta imunológica após um conjunto de eventos denominado de apresentação de antígeno
- Para a resposta imunológica adquirida, o organismo entra em contato com um elemento estranho pela primeira vez, sendo este fagocitado e processado por células apresentadoras de antígeno que vão ativar células T auxiliares. Essas células vão estimular as células B e T a se diferenciarem e formarem células comprometidas com o reconhecimento, ataque e destruição dos antígenos. As células B vão formar os plasmócitos, que produzirão imunoglobulinas específicas, resultando na resposta imunológica humoral. As células T formarão as células T citotóxicas, originando a resposta imunológica celular
- As moléculas localizadas na superfície das células participantes da formação dos complexos de apresentação de antígenos são chamadas de proteínas do complexo principal de histocompatibilidade, ou MHC, que é um conjunto de genes que codificam proteínas de superfície celular e que formarão um complexo proteico com o antígeno processado. Essas proteínas também fornecem identidade às células de um mesmo indivíduo, reconhecidas como próprias do organismo

- As imunoglobulinas constituem os anticorpos que têm estrutura proteica em forma de "Y", constituído por duas cadeias pesadas e duas cadeias leves, similares, ligadas por pontes dissulfeto. Essas moléculas apresentam uma porção constante, na região C-terminal e porções variáveis, localizadas na região N-terminal, responsáveis pela diversidade e pela especificidade da resposta imunológica
- A diversidade estrutural dos anticorpos, assim como dos receptores celulares específicos, acontecerá por mecanismos genéticos de rearranjos por recombinação, resultando em alinhamento de segmentos gênicos e formando sequências adequadas para a expressão das proteínas da resposta imunológica, durante a diferenciação dos linfócitos B e T
- As alergias e as doenças autoimunes são exemplos de disfunção da resposta imunológica. Já a vacinação é um exemplo de utilização dessa resposta como maneira de proteção de indivíduos contra doenças contagiosas.

AUTOAVALIAÇÃO

7.1 Na sua opinião, como o estudo do genoma humano pode influenciar a descoberta de patologias humanas?
7.2 O que são erros inatos do metabolismo?
7.3 Como você explica o papel das enzimas no desenvolvimento de doenças genéticas do metabolismo? Dê exemplos.
7.4 Quais são as principais doenças descritas associadas ao metabolismo do aminoácido fenilalanina?
7.5 Para você, como crianças portadoras de mutações em genes de metabolização do glicídio galactose devem ser tratadas?
7.6 Como mutações nos receptores de lipoproteínas de baixa densidade ou LDL podem causar hipercolesterolemia familiar?
7.7 Defina porfirias.
7.8 O que é terapia gênica?
7.9 Descreva, em linhas gerais, a estrutura da hemoglobina.
7.10 Diferencie a hemoglobina adulta da hemoglobina fetal.
7.11 O que são as regiões controladoras de *locus* dos genes das globinas?
7.12 Como se dividem as hemoglobinopatias?
7.13 Como você definiria o estado de um indivíduo portador do chamado traço falcêmico?
7.14 O que são talassemias?
7.15 Diferencie imunidade inata de imunidade adquirida.
7.16 Conceitue antígeno e determinante antigênico.
7.17 O que são células apresentadoras de antígenos?
7.18 Como a expressão de antígenos HLA auxiliam no reconhecimento do que é próprio do organismo?
7.19 Quais são as principais classes de anticorpos produzidas pelos vertebrados?
7.20 Qual é o tipo de mecanismo genético envolvido na produção de anticorpos específicos para os diferentes determinantes antigênicos os quais estamos expostos?
7.21 Defina alergia e doença autoimune.

BIBLIOGRAFIA

ALLISON, A.C. Protection afforded by sickle cell trait against subtertian malarial infection. *British Medical Journal* 1954; Feb: 290-294.

BORDIGNON, C.; NOTARANGELO, L.D.; NOBILI, N. *et al.* Gene therapy in peripheral blood lymphocytes and bone marrow for ADA-immunodeficient patients. *Science* 1995 October; 20270: 470-475.

CAMPBELL, N.A.; REECE, J.B. *Biologia.* 8ª ed., Porto Alegre: Artmed, 2010.

DEVLIN, T.M. *Textbook of biochemistry – with clinical correlation.* 4ª edição, Wiley-Liss, EUA, 1997.

NUSSBAUM, R.L.; MCINNES, R.R.; WILLARD, H.F. *Thompson & Thompson genética médica.* 7ª ed., Rio de Janeiro: Elsevier, 2007.

OMIM® – Online Mendelian Inheritance in Man®. Disponível em: http://www.ncbi.nlm.nih.gov/omim e http://globina.cse.psu.edu/. Acesso em 01/02/2012.

SAVITT, T.L.; GOLDBERG, M.F. Herrick's 1910 case report of sickle cell anemia. The rest of the story. *Jama.* 1989; Jan 13;261(2):266-71.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

STRACHAN, T.; READ, A.P. *Genética molecular humana.* 2ª ed., Porto Alegre: Artmed, 2002.

TURNPENNY, P.D.; ELLARD, S. *Emery genética médica.* 13ª ed., Rio de Janeiro: Elsevier, 2009.

YOUNG, I.D. *Genética médica.* Rio de Janeiro: Guanabara Koogan, 2007.

8

Genética de Populações

Objetivos de estudo

Compreender a dinâmica dos genes nas populações

Saber calcular as frequências gênicas e genotípicas com base nos dados populacionais, tanto para alelos autossômicos quanto para genes ligados ao X

Compreender o modelo de Hardy-Weinberg e reconhecer as condições para sua utilização

Ser capaz de testar o equilíbrio de Hardy-Weinberg por meio da aplicação do qui-quadrado

Conhecer os tipos de cruzamento existentes e os efeitos da endogamia

Saber quais agentes são capazes de promover variações genéticas nas populações e como cada um deles atua

Conhecer os principais componentes que influenciam a adaptabilidade de um traço

Reconhecer de que maneira o ambiente atua, moldando a variação fenotípica nas populações

Compreender como o fluxo gênico pode interferir na composição genética das populações envolvidas

Saber o que representa valor adaptativo e coeficiente de seleção

Perceber como diferentes alelos são mantidos em uma população como polimorfismo balanceado

Ser capaz de conceituar e distinguir os tipos de seleção que atuam em traços quantitativos: estabilizadora, direcional e disruptiva

Conhecer os efeitos que a deriva genética aleatória pode causar nas frequências gênicas e genotípicas em demes pequenos

Saber definir efeito fundador e gargalo genético

Conceitos-chave do capítulo

Adaptabilidade
Autofecundação
Codominante (codominância)
Coeficiente de endocruzamento
Coeficiente de seleção (*s*)
Crossing over
Cruzamento aleatório
Deme
Deriva genética
Efeito fundador
Endocruzamento (endogamia)
Equilíbrio de Hardy-Weinberg
Erro de amostragem
Fluxo gênico
Frequência alélica
Frequência gênica
Frequência genotípica
Gargalo genético
Grau de liberdade (*gl*)
Hemizigoto (hemizigose)
Lei de Hardy-Weinberg
Migração
Mutação
Mutação direta
Mutação reversa
Panmixia
Penetrância completa
Pool gênico
População
Qui-quadrado
Recombinação gênica
Seleção balanceadora
Seleção estabilizadora
Seleção direcional
Seleção disruptiva
Sexo heterogamético
Sexo homogamético
Sobredominância
Valor adaptativo (*W*)
Vantagem do heterozigoto

Introdução

A capacidade de duplicação semiconservativa da molécula de DNA constitui a base da unidade e da continuidade genética entre os seres vivos ao longo do tempo. Neste contexto, e considerando que os organismos naturalmente vivem em grupos, além da análise da transmissão gênica nas famílias, já abordada, um enfoque igualmente importante refere-se à *dinâmica dos genes nas populações*. Cada variante alélica em uma população sexuada, em qualquer época, é descendente de um gene que existiu no passado, significando que a composição genética da população detém fonte importante de informação sobre a sua trajetória evolutiva.

A genética de populações trata da aplicação dos conceitos genéticos no contexto populacional, buscando, entre outras questões, compreender o significado biológico e evolutivo da

variação genética. Por diferentes abordagens, é possível avaliar a distribuição e a transmissão de alelos e genótipos nas populações, definir suas frequências e os padrões de variação genética entre grupos de organismos que entrecruzam, bem como compreender as divergências genéticas entre as espécies. Outro foco da genética de populações refere-se ao estudo das *forças evolutivas* (*mutação, migração, seleção natural* e *deriva genética*) capazes de alterar as *frequências gênicas* e *genotípicas* nas populações pelas gerações. Fundamentalmente, o processo evolutivo envolve mudanças na estrutura genética populacional ao longo do tempo. Assim, para traçar a história evolutiva de uma população, convém conhecer sua composição genética, sendo inúmeras as aplicações desse conhecimento em áreas como a medicina e o melhoramento animal e vegetal.

As populações são as unidades básicas da evolução, que ocorre em virtude de mudanças genéticas. Pela dinâmica dos genes nas populações e das forças que moldam a variação genética, entende-se o processo evolutivo.

■ Distribuição de genes e genótipos nas populações

Na natureza, são raros os seres que sobrevivem de maneira isolada. Em geral, organismos sexuados vivem em grupos, adaptados a um ambiente particular. Na genética de populações, o termo população aplica-se a um grupo de organismos da mesma espécie que ocupam determinado espaço geográfico e são capazes de acasalar entre si. Ainda que, por definição, o entrecruzamento seja inerente a todos os seres sexuados de uma espécie, essencialmente o acasalamento não ocorre de maneira indiscriminada. Vejamos o exemplo de insetos que habitam determinada floresta. Embora aqueles de uma mesma espécie possam cruzar entre si, em geral, o acasalamento se dá entre os que habitam uma mesma árvore, raramente ocorrendo entre insetos que habitam árvores distantes. Percebemos, portanto, que a diversidade ambiental, incluindo as inúmeras barreiras geográficas existentes, é um dos fatores que interfere no acasalamento, especificamente nas possibilidades de escolha dos parceiros. Além do ambiente, outras circunstâncias também podem influenciar o critério de escolha, como fatores culturais, religiosos ou socioeconômicos na espécie humana. Desse modo, a troca de genes ocorre, efetivamente, pelo entrecruzamento de organismos em *unidades populacionais locais*, também denominadas *subpopulações* ou demes. Ainda que cada ser vivo seja uma combinação gênica única, os diversos organismos que constituem essas *populações locais de entrecruzamento* partilham um *estoque genético* representado pelo conjunto de alelos existentes (*pool* gênico), que lhes confere identidade genética. Na prática, é comum utilizarmos o termo *população* em referência às *populações locais de entrecruzamento*.

População
Grupo de organismos de uma mesma espécie, que vivem na mesma área geográfica, partilham um conjunto gênico comum e são capazes de cruzar entre si e produzir prole fértil

Deme
População local de organismos que entrecruzam

***Pool* gênico**
Conjunto de todos os alelos compartilhados pelos organismos que constituem uma população local de entrecruzamento em determinada época

Cada unidade populacional (deme) compartilha um mesmo conjunto gênico.

Geração após geração, pela segregação e pela recombinação alélica, os genótipos dos organismos que compõem uma população sexuada dão origem a novos gametas, os quais, por fertilização, formam novas combinações de genótipos nos zigotos. Desse modo, elos reprodutivos conferem à população uma continuidade no tempo e no espaço, uma vez que os genótipos existentes em uma geração de organismos reprodutores determinarão os genótipos da geração seguinte. Como mostrado na Figura 8.1, os genótipos presentes em determinada geração são resultantes da combinação de alelos produzidos na geração anterior e, por sua vez, irão segregar e recombinar para formar os gametas que darão origem aos genótipos da geração seguinte, embora eles mesmos não sejam transmitidos de maneira direta pelas gerações. Assim,

Figura 8.1 Interconexão reprodutiva existente entre as gerações de uma população.

a cada geração, o conjunto gênico da população é recomposto, de maneira que toda população pode ser caracterizada pela distribuição de seus genes.

Frequências gênicas e genotípicas | Como calcular?

Frequência alélica (ou gênica)
Refere-se à proporção dos alelos de um gene em uma população

Frequência genotípica
Refere-se à proporção de cada genótipo em um *locus* específico entre os membros de uma população

Denomina-se frequência alélica (ou gênica) a proporção de cada alelo presente em determinado *locus* em uma população e frequência genotípica à proporção dos genótipos resultantes da combinação desses alelos.

Em grandes populações, em geral, não é possível proceder à análise genética em todos os organismos, sendo os cálculos das *frequências alélicas* e *genotípicas* feitos de amostras de organismos representativas dessas populações. As *frequências genotípicas* são obtidas dividindo-se o número de indivíduos com determinado genótipo pelo número total de indivíduos na amostra analisada (N). Considerando a situação simples de um *locus* autossômico com um par de alelos (A e a), sem dominância, e os 3 genótipos possíveis (AA, Aa e aa), a frequência (f) de cada genótipo equivale a:

$$f(AA) = \frac{\text{Número de indivíduos } AA}{N}$$
$$f(Aa) = \frac{\text{Número de indivíduos } Aa}{N} \qquad \text{(Equação 8.1)}$$
$$f(aa) = \frac{\text{Número de indivíduos } aa}{N}$$

A soma das frequências genotípicas é igual a 1.

Como vimos, em populações de organismos sexuados, os genótipos são combinações transitórias de alelos que, a cada geração, se desfazem. Os alelos é que são transmitidos para a geração seguinte pelos gametas. Desse modo, a população pode ser caracterizada por sua *distribuição gênica*, ou seja, pelas frequências dos alelos existentes nos diferentes *loci*.

A frequência de determinado alelo em uma população é expressa como o número de cópias desse gene dividido pelo número total de alelos no *locus*. Logo, em uma população diploide, as *frequências gênicas* de um par de alelos autossômicos (A e a), representadas por p e q, respectivamente, podem ser obtidas da seguinte maneira:

$$p = f(A) = \frac{(2n_{AA} + n_{Aa})}{2N} \qquad \text{(Equação 8.2)}$$
$$q = f(a) = \frac{(2n_{aa} + n_{Aa})}{2N}$$

sendo: n_{AA}, n_{Aa} e n_{aa} = número de indivíduos com os genótipos AA, Aa e aa; N = número total de indivíduos.

Verificamos, portanto, que a frequência de determinado alelo (A ou a) corresponde a duas vezes o número de indivíduos homozigotos para o alelo em questão, acrescido do número de heterozigotos e dividido por $2N$, pois os organismos diploides têm um par de alelos em cada *locus*. A soma das *frequências gênicas* sempre é igual a 1, ou seja, $p + q = 1$.

É possível, também, calcular a proporção de cada alelo conhecendo-se as frequências dos genótipos existentes na população em estudo, isto é, as *frequências gênicas* podem ser calculadas com base nas *frequências genotípicas*, como mostrado a seguir:

$$p = f(A) = f(AA) + ½\ f(Aa) \qquad \text{(Equação 8.3)}$$

$$q = f(a) = f(aa) + ½\ f(Aa)$$

A frequência de determinado alelo em um *locus* equivale à soma da frequência do genótipo homozigoto para o alelo com metade da frequência dos heterozigotos, os quais têm apenas uma cópia do alelo. Vejamos como efetuar esse cálculo, retomando o exemplo da Tabela 8.1,

Codominante
Quando os efeitos de 2 alelos distintos presentes em um *locus* no par de cromossomos homólogos são observados no fenótipo

referente ao grupo sanguíneo humano MN. Nesse *locus*, verifica-se a existência de 2 alelos codominantes L^M e L^N, cuja combinação resulta em 3 genótipos possíveis (L^ML^M, L^ML^N e L^NL^N), os quais, por sua vez, correspondem a 3 fenótipos distintos (M, MN e N). Com base na determinação dos antígenos M e N nas hemácias de 1.200 indivíduos da população norte-americana, foram identificados 420 L^ML^M, 600 L^ML^N e 180 L^NL^N. Se dividirmos o número de indivíduos com cada genótipo pelo número total de pessoas analisadas, obteremos as frequências dos 3 genótipos:

$$\text{Frequência do genótipo } L^ML^M = 420/1.200 = 0{,}35$$
$$\text{Frequência do genótipo } L^ML^N = 600/1.200 = 0{,}50$$
$$\text{Frequência do genótipo } L^NL^N = 180/1.200 = 0{,}15$$

Os cálculos mostram que 50% dos indivíduos dessa população são heterozigotos, 35% são homozigotos L^ML^M e 15% são homozigotos L^NL^N, sendo a soma das frequências genotípicas igual a 1 ou 100%. Conhecendo-se as frequências dos diferentes genótipos, podemos calcular as *frequências gênicas* ao aplicar a Equação 8.3:

$$\text{Frequência do alelo } L^M = p = (0{,}35) + ½\ 0{,}50 = 0{,}60$$
$$\text{Frequência do alelo } L^N = q = (0{,}15) + ½\ 0{,}50 = 0{,}40$$

$$p + q = 0{,}60 + 0{,}40 = 1$$

Vimos, portanto, que a frequência do alelo L^M (p) foi calculada somando-se a frequência do genótipo homozigoto L^ML^M com metade da frequência dos heterozigotos; de maneira análoga, a frequência do alelo L^N (q) foi calculada somando-se a porcentagem do genótipo homozigoto L^NL^N com metade da frequência dos heterozigotos.

As *frequências gênicas* constituem um dos indicadores da diversidade genética existente entre as populações, de maneira que é comum encontrarmos variações nas *frequências alélicas* em diferentes *demes*. A título de ilustração, na amostra de indivíduos norte-americanos do exemplo, as frequências dos alelos L^M e L^N (0,60 e 0,40, respectivamente) se apresentam bem distintas das frequências desses mesmos alelos entre aborígines australianos (0,18 e 0,82, respectivamente) e esquimós (0,91 e 0,09, respectivamente).

Populações que apresentam *frequências gênicas* iguais não necessariamente terão as mesmas *frequências genotípicas*, ou seja, *demes* distintos podem ter *frequências gênicas* iguais e diferir em relação às frequências de seus genótipos. A Tabela 8.1 apresenta dados referentes ao grupo sanguíneo MN, que nos possibilitam comparar as *frequências gênicas* e *genotípicas* em 3 populações diferentes.

■ **Tabela 8.1** Frequências genotípicas e gênicas relativas aos alelos do grupo sanguíneo MN observadas em 3 demes distintos.

	Frequências genotípicas			Frequências gênicas	
	L^ML^M	L^ML^N	L^NL^N	p	q
População I	0,35	0,50	0,15	0,6	0,4
População II	0,50	0,20	0,30	0,6	0,4
População III	0,36	0,48	0,16	0,6	0,4

Observe que, apesar de as frequências dos alelos L^M e L^N serem iguais nas 3 populações (p = 0,60 e q = 0,40, respectivamente), as *frequências genotípicas* são nitidamente distintas.

Loci com alelos múltiplos

A existência de mais de 2 alelos é comum em muitos *loci* e, nesses casos, os mesmos princípios já apresentados podem ser usados no cálculo das *frequências gênicas*. Vejamos a situação de 3 alelos (A_1, A_2 e A_3) em um *locus*, cuja combinação produz seis genótipos (A_1A_1,

A_1A_2, A_1A_3, A_2A_2, A_2A_3 e A_3A_3). As *frequências alélicas* podem ser calculadas com base nas *frequências genotípicas*, aplicando-se a mesma lógica da Equação 8.3:

$$p = f(A_1) = f(A_1A_1) + ½\, f(A_1A_2) + ½\, f(A_1A_3)$$

$$q = f(A_2) = f(A_2A_2) + ½\, f(A_1A_2) + ½\, f(A_2A_3) \qquad \text{(Equação 8.4)}$$

$$r = f(A_3) = f(A_3A_3) + ½\, f(A_1A_3) + ½\, f(A_2A_3)$$

Sexo heterogamético
Aquele que, em relação aos cromossomos sexuais, produz 2 tipos de gametas. Por exemplo, em mamíferos, o macho é o sexo heterogamético porque produz gametas com o cromossomo X e gametas com o cromossomo Y

Hemizigoto (hemizigose)
Aquele que tem somente um alelo em determinado *locus* nas células diploides. Os homens são hemizigotos para genes ligados ao X por terem apenas um cromossomo X

Sexo homogamético
Aquele que, em relação aos cromossomos sexuais, produz apenas um tipo de gameta. Por exemplo, as fêmeas de mamíferos representam o sexo homogamético, pois todos os gametas produzidos contêm o cromossomo X

Loci ligados ao cromossomo X

Em relação aos genes localizados no X, os princípios mostrados também são válidos para o cálculo das *frequências gênicas*, lembrando, no entanto, que no sistema de determinação sexual XY, como o que ocorre em mamíferos, o sexo heterogamético (XY) é hemizigoto para alelos do cromossomo X, ao passo que o sexo homogamético (XX) pode ser homozigoto ou heterozigoto. Por essa razão, alelos ligados ao X não se distribuem igualmente na população, uma vez que ocorrem em desproporção nos 2 sexos: 2/3 deles estão no *sexo homogamético* e 1/3 no *sexo heterogamético*. Supondo-se a existência de 2 alelos em um *locus* do cromossomo X humano (X_A e X_a), as mulheres podem apresentar 3 genótipos (X_AX_A, X_AX_a e X_aX_a), ao passo que os homens, apenas 2, X_AY ou X_aY. Desse modo, em uma população com igual número de indivíduos masculinos e femininos, o cálculo das *frequências alélicas* no *conjunto gênico total* da população deverá levar em conta as frequências dos diferentes genótipos em cada sexo:

$$p = f(X_A) = \frac{f(X_AY) + f(X_AX_a) + 2\, f(X_AX_A)}{3} \qquad \text{(Equação 8.5)}$$

$$q = f(X_a) = \frac{f(X_aY) + f(X_AX_a) + 2\, f(X_aX_a)}{3} \quad \text{ou} \quad q = 1 - p$$

A composição genética de um *deme*, representada por um conjunto gênico próprio, pode ser definida em termos de suas *frequências gênicas* e *genotípicas*.

Modelo de Hardy-Weinberg

Da caracterização de determinada população em termos de suas *frequências gênicas* e *genotípicas* e, sabendo que as diferentes gerações dessa população estão unidas por elos de reprodução, com consequente transmissão gênica, como determinar as *frequências alélicas* e *genotípicas*, sucessivamente, a cada geração?

O *princípio de Hardy-Weinberg*, formulado em 1908 de maneira independente pelo matemático inglês Godfrey Harold Hardy e pelo médico alemão Wilhelm Weinberg, estabelece uma relação matemática entre as *frequências gênicas* e *genotípicas* em gerações subsequentes (reprodutores e seus descendentes). Esse *modelo de equilíbrio genético*, embasado nas leis mendelianas de segregação, nos possibilita fazer algumas predições sobre as frequências dos alelos e suas combinações nos genótipos, sendo seu uso válido, desde que algumas condições sejam atendidas. Trata-se de um modelo de referência muito utilizado em estudos populacionais que assume a não interferência de *forças evolutivas* na população e que serve de base na comparação com outros modelos matemáticos mais complexos, os quais, por sua vez, não descartam a atuação de diferentes agentes e seus efeitos na sobrevivência diferencial dos genótipos e na viabilidade dos gametas.

Equilíbrio de Hardy-Weinberg
Condição de um estado de equilíbrio genético, na manutenção das frequências alélicas e genotípicas de uma geração para outra em grandes populações com reprodução aleatória, que não sofrem a ação de mutação, migração ou seleção natural. Também chamada lei, princípio ou modelo de Hardy-Weinberg

As predições do equilíbrio de Hardy-Weinberg podem ser aplicadas a grandes populações de organismos diploides, com reprodução sexuada e *cruzamento aleatório* (em relação ao traço em estudo), nas quais os casais sejam igualmente férteis e as *frequências alélicas* sejam as mesmas nos machos e nas fêmeas. Além disso, assume-se que *forças evolutivas* sistemáticas, com potencial de alterar as *frequências gênicas* na população, não estão atuando ou são desprezíveis.

Entre as principais condições necessárias à utilização desse modelo destaca-se a necessidade da população em estudo ser grande o suficiente, de maneira que as *frequências alélicas*

não sofram variações em virtude de *erros de amostragem*, como veremos mais adiante neste capítulo, ao abordarmos os efeitos da *deriva genética aleatória*. Considera-se também que as gerações sejam discretas, ou seja, não ocorra sobreposição entre elas. Além disso, outro fator determinante das *frequências genotípicas* é a maneira como os casais se constituem a cada geração, ou seja, o tipo de acasalamento existente entre os organismos da população. O *princípio de Hardy-Weinberg* baseia-se no cruzamento aleatório (ou panmixia) que representa um sistema de reprodução comumente observado nas populações. Sob *panmixia*, a probabilidade de 2 organismos com genótipos específicos se cruzarem será igual ao produto das frequências de seus genótipos na população. Portanto, em uma *população panmítica*, na qual os acasalamentos são ao acaso, a probabilidade de 2 genótipos colidirem será determinada por suas respectivas frequências na população.

Cruzamento aleatório (ou panmixia)
Reprodução aleatória entre os organismos de uma população. Corresponde à associação ao acaso de genótipos

Em uma *unidade panmítica*, um indivíduo tem igual probabilidade de se acasalar com qualquer outro do sexo oposto, independentemente de seu fenótipo.

Outra condição importante à aplicação do *modelo de Hardy-Weinberg* refere-se à ausência, na população, da ação de *forças evolutivas*, como *mutação*, *migração* e *seleção natural*, que podem promover alterações nas frequências dos alelos ao longo das gerações. Considerando que a reprodução sexuada, por si, não é capaz de alterar as *frequências gênicas* ou *genotípicas* e, em se tratando de uma população grande e *panmítica*, com gerações discretas, que atenda aos demais pressupostos citados, o *princípio de Hardy-Weinberg* estabelece algumas previsões importantes:

- As frequências dos genótipos (*frequências genotípicas*) podem ser calculadas com base nas frequências dos alelos em um *locus* específico (*frequências gênicas*)
- Na ausência de *forças evolutivas* sistemáticas, as *frequências gênicas* não se alteram de uma geração para a seguinte em grandes populações *panmíticas*;
- As frequências genotípicas alcançam o equilíbrio de Hardy-Weinberg e se mantêm constantes após uma geração de cruzamentos aleatórios.

Desse modo, se uma população cumprir os pressupostos estabelecidos pelo *modelo de Hardy-Weinberg* (Tabela 8.2) ela manterá a estabilidade genética, de maneira que suas *frequências gênicas* e *genotípicas* permanecerão inalteradas pelas gerações. É importante perceber a grande utilidade desse modelo no campo da genética de populações, pois, à medida que o *princípio de Hardy-Weinberg* aponta as condições que mantêm as frequências de alelos e genótipos constantes, ele também nos possibilita isolar e quantificar as *forças evolutivas* que estão atuando na população. Deve-se ressaltar que esse modelo avalia a estabilidade das *frequências genotípicas* em *loci* individuais, o que significa que uma população pode estar em equilíbrio em relação aos alelos de determinado *locus*, mas não em relação a outros.

O *modelo de Hardy-Weinberg* nos possibilita calcular as *frequências genotípicas* com base nas *frequências gênicas* (e vice-versa) em populações grandes e *panmíticas*.

■ **Tabela 8.2** Principais condições necessárias à aplicação da *lei de Hardy-Weinberg*.

População, infinitamente grande, de organismos diploides
Reprodução sexuada e aleatória (panmixia). Não ocorrência de cruzamento seletivo ou endocruzamento. O cruzamento ao acaso de genótipos equivale à união aleatória de gametas
Gerações não sobrepostas
Todos os genótipos (zigotos e adultos), bem como os gametas, são igualmente viáveis e férteis, o que representa ausência de seleção natural
As frequências dos alelos são iguais em ambos os sexos
Inexistência de novos alelos criados por mutação
Ausência de migração (fluxo gênico) e deriva genética aleatória

Equação de Hardy-Weinberg

Se considerarmos a união aleatória de gametas femininos e masculinos igualmente viáveis, produzidos em determinada geração de uma população, poderemos, com base nas *frequências gênicas* nesse conjunto de gametas, calcular as frequências dos genótipos que se formam nos zigotos, assumindo que todos os indivíduos da população (machos e fêmeas) contribuem igualmente para o reservatório de gametas. Desse modo, em se tratando de um *locus* autossômico com 2 alelos, como o do grupo sanguíneo MN, o conjunto de gametas (masculinos e femininos) é representado por uma fração de gametas com o alelo L^M (p) e uma fração de gametas com o alelo L^N (q), como mostrado na Figura 8.2.

♂ \ ♀	L^M p	L^N q
L^M p	L^ML^M p^2	L^ML^N pq
L^N q	L^ML^N pq	L^NL^N q^2

Figura 8.2 Proporção de genótipos produzidos pela combinação de 2 alelos em um *locus* autossômico com base na união aleatória de gametas masculinos e femininos.

A combinação de alelos resultante da união aleatória dos gametas masculinos e femininos origina zigotos L^ML^M, L^ML^N e L^NL^N na seguinte proporção: $(p_♂ + q_♂)$ $(p_♀ + q_♀)$. Assumindo que as frequências dos alelos L^M e L^N são iguais nos gametas de ambos os sexos, ou seja, $p_♀ = p_♂$ e $q_♀ = q_♂$, verificamos que a distribuição das *frequências genotípicas* nos zigotos é equivalente a:

$$(p + q)\ (p + q) = (p + q)^2 = p^2 + 2pq + q^2 \qquad \text{(Equação 8.6)}$$

p e q correspondem às frequências dos alelos L^M e L^N, respectivamente.

p^2, $2pq$ e q^2 correspondem às frequências dos genótipos L^ML^M, L^ML^N e L^NL^N, respectivamente.

Revendo a Figura 8.2, é fácil verificar que a probabilidade de um espermatozoide e um ovócito, portadores do alelo L^M, se encontrarem é $p \times p = p^2$; a chance de 2 gametas com o alelo L^N se unirem é $q \times q = q^2$ e a frequência de gametas masculinos portadores do alelo L^M se juntarem a gametas femininos com o alelo L^N, ou vice-versa, é equivalente a $p \times q + q \times p = 2pq$. Concluímos, portanto, que o *cruzamento aleatório* produz genótipos nas proporções p^2, $2pq$ e q^2.

A Equação 8.6, conhecida como *equação de Hardy-Weinberg*, torna clara a nítida relação existente entre as *frequências gênicas* nos gametas e as *frequências genotípicas* nos zigotos, sendo possível calcular as *frequências genotípicas* no equilíbrio com base nas *frequências alélicas*. Essa relação pode ser estendida para *loci* com qualquer número de alelos (*alelos múltiplos*) e maior detalhamento sobre o assunto é dado na seção Alelos autossômicos múltiplos.

O *modelo de Hardy-Weinberg* estabelece, portanto, que uma população cujas *frequências genotípicas* se apresentam de acordo com a Equação 8.6 está em *equilíbrio genético*, sendo este alcançado em apenas uma geração de *cruzamentos aleatórios*. Quando em *equilíbrio de Hardy-Weinberg*, as *frequências gênicas* e *genotípicas* se mantêm constantes no decorrer das gerações e somente sofrerão alterações quando *forças evolutivas* atuarem.

Para compreender a *lei de Hardy-Weinberg*, tomemos por base uma população na qual os indivíduos entrecruzam de maneira aleatória, de maneira que, a chance de reprodução de cada genótipo está diretamente relacionada a sua frequência. A Tabela 8.3 apresenta a distribuição em uma população dos tipos de cruzamentos possíveis relativos aos genótipos formados pela combinação de um par de alelos autossômicos, L^M e L^N, com frequências p e q, bem como, a proporção dos diferentes genótipos entre os descendentes.

■ **Tabela 8.3** Demonstração da *lei de Hardy-Weinberg* para um *locus* autossômico com um par de alelos.

Casais (geração inicial)		Frequências de zigotos (descendentes)		
Cruzamento	Frequência	L^ML^M	L^ML^N	L^NL^N
$L^ML^M \times L^ML^M$	$p^2 \times p^2 = p^4$	p^4	–	–
$L^ML^M \times L^ML^N$	$2 \times p^2 \times 2pq = 4p^3q$	$2p^3q$	$2p^3q$	–
$L^ML^M \times L^NL^N$	$2 \times p^2 \times q^2 = 2p^2q^2$	–	$2p^2q^2$	–
$L^ML^N \times L^ML^N$	$2pq \times 2pq = 4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
$L^ML^N \times L^NL^N$	$2 \times 2pq \times q^2 = 4pq^3$	–	$2pq^3$	$2pq^3$
$L^NL^N \times L^NL^N$	$q^2 \times q^2 = q^4$	–	–	q^4
Total	$(p + q)^4 = 1$	p^2	$2pq$	q^2

Sendo $p + q = 1$, a soma das frequências dos diferentes tipos de cruzamentos em uma geração de uma *população panmítica*, equivalente a $(p + q)^4$, também será igual a 1. Somando-se as colunas das frequências dos zigotos com os genótipos L^ML^M, L^ML^N e L^NL^N, veremos que a distribuição desses genótipos nessa primeira geração de descendentes se apresenta de acordo com a equação $p^2 + 2pq + q^2$, assim:

$$L^ML^M = p^4 + 2p^3q + p^2q^2 = p^2 (p^2 + 2pq + q^2) = p^2$$
$$L^ML^N = 2p^3q + 4p^2q^2 + 2pq^3 = 2pq (p^2 + 2pq + q^2) = 2pq$$
$$L^NL^N = p^2q^2 + 2pq^3 + q^4 = q^2 (p^2 + 2pq + q^2) = q^2$$

Observe que as frequências dos genótipos L^ML^M e L^NL^N na geração de descendentes correspondem ao quadrado da frequência dos alelos L^M e L^N, expressas como p^2 e q^2, respectivamente. A frequência do genótipo L^ML^N entre os descendentes, por sua vez, corresponde ao dobro do produto das frequências de cada alelo ($2pq$).

Vejamos, a seguir, a extensão desse conhecimento usando os dados numéricos da amostra dos 1.200 indivíduos norte-americanos testados para o grupo sanguíneo MN, na qual os genótipos se distribuem nas seguintes frequências: $L^ML^M = 0,35$, $L^ML^N = 0,50$ e $L^NL^N = 0,15$, sendo as frequências dos alelos L^M e L^N iguais a 0,60 (p) e 0,40 (q), respectivamente. Todos os genótipos conferem a seus portadores taxas iguais de sobrevivência e de reprodução. Consideraremos essa a geração inicial e, assumindo tratar-se de uma população *panmítica*, que obedece aos pressupostos do *modelo de Hardy-Weinberg*, as frequências das combinações possíveis de casais para os genótipos L^ML^M, L^ML^N e L^NL^N se dará como especificado na segunda coluna da Tabela 8.4A.

■ **Tabela 8.4A** Demonstração da *lei de Hardy-Weinberg* para o *locus* autossômico do grupo sanguíneo MN, sendo as frequências dos genótipos L^ML^M, L^ML^N e L^NL^N na geração inicial iguais a 0,35, 0,50 e 0,15, respectivamente. Distribuição dos genótipos L^ML^M, L^ML^N e L^NL^N na primeira geração de descendentes, obtida com base no cruzamento aleatório de indivíduos da geração inicial.

Casais (geração inicial)		Frequências de zigotos (1ª geração de descendentes)		
Cruzamento	Frequência	L^ML^M	L^ML^N	L^NL^N
$L^ML^M \times L^ML^M$	$0,35 \times 0,35 = 0,1225$	0,1225	–	–
$L^ML^M \times L^ML^N$	$2 \times 0,35 \times 0,50 = 0,35$	0,175	0,175	–
$L^ML^M \times L^NL^N$	$2 \times 0,35 \times 0,15 = 0,105$	–	0,105	–
$L^ML^N \times L^ML^N$	$0,50 \times 0,50 = 0,25$	0,0625	0,125	0,0625
$L^ML^N \times L^NL^N$	$2 \times 0,50 \times 0,15 = 0,15$	–	0,075	0,075
$L^NL^N \times L^NL^N$	$0,15 \times 0,15 = 0,0225$	–	–	0,0225
Total	1,0000	0,3600	0,4800	0,1600

■ **Tabela 8.4B** Distribuição dos genótipos L^ML^M, L^ML^N e L^NL^N na segunda geração de descendentes, obtida do cruzamento aleatório de indivíduos da primeira geração de descendentes.

Casais (1ª geração de descendentes)		Frequências de zigotos (2ª geração de descendentes)		
Cruzamento	Frequência	L^ML^M	L^ML^N	L^NL^N
$L^ML^M \times L^ML^M$	$0,36 \times 0,36 = 0,1296$	0,1296	–	–
$L^ML^M \times L^ML^N$	$2 \times 0,36 \times 0,48 = 0,3456$	0,1728	0,1728	–
$L^ML^M \times L^NL^N$	$2 \times 0,36 \times 0,16 = 0,1152$	–	0,1152	–
$L^ML^N \times L^ML^N$	$0,48 \times 0,48 = 0,2304$	0,0576	0,1152	0,0576
$L^ML^N \times L^NL^N$	$2 \times 0,48 \times 0,16 = 0,1536$	–	0,0768	0,0768
$L^NL^N \times L^NL^N$	$0,16 \times 0,16 = 0,0256$	–	–	0,0256
Total	1,0000	0,3600	0,4800	0,1600

Sendo os casais igualmente férteis, as *frequências genotípicas* esperadas entre seus descendentes, segundo as leis de Mendel, são aquelas apresentadas nas 3 últimas colunas da Tabela 8.4A, totalizando 36% de indivíduos com o genótipo L^ML^M, 48% com o genótipo L^ML^N e 16% com o genótipo L^NL^N na primeira geração de zigotos. Assim, as *frequências genotípicas* dessa primeira geração de descendentes (0,36 L^ML^M: 0,48 L^ML^N: 0,16 L^NL^N) equivalem às *proporções de Hardy-Weinberg* (p^2: $2pq$: q^2), como demonstrado na Tabela 8.3. A partir daí, podemos calcular as frequências dos alelos L^M e L^N nessa geração de descendentes. Aplicando-se a Equação 8.3, obtemos valores 0,60 (p) e 0,40 (q). Portanto, os dados da Tabela 8.4A evidenciam que as *frequências gênicas* (p e q) são iguais no *conjunto gênico* da geração inicial e naquele da primeira geração de descendentes, o que nos possibilita concluir a validade ao *modelo de Hardy-Weinberg* em relação ao fato de, sob condições de equilíbrio, as *frequências alélicas* não se alterarem ao longo das gerações.

Note, no entanto, que embora as *frequências gênicas* (p e q) se mantenham constantes no decorrer das gerações (geração inicial e primeira geração de descendentes), o mesmo não ocorreu com as *frequências genotípicas* (L^ML^M, L^ML^N e L^NL^N), representadas, respectivamente, por 0,35, 0,50 e 0,15 na geração inicial e por 0,36, 0,48 e 0,16 na primeira geração de zigotos. A validade da *lei de Hardy-Weinberg*, no que se refere às *frequências genotípicas* da população alcançarem o equilíbrio e se manterem estáveis após uma geração de *panmixia*, pode ser verificada na Tabela 8.4B, que nos mostra o que ocorre quando a primeira geração de descendentes amadurece e se reproduz. Vejamos qual será a nova distribuição de frequências na geração seguinte. Observe que os genótipos L^ML^M, L^ML^N e L^NL^N na segunda geração de descendentes se apresentam, respectivamente, nas proporções 0,36, 0,48 e 0,16, equivalentes a p^2, $2pq$ e q^2, como demonstrado na Tabela 8.3, e que essas *frequências genotípicas* são iguais às da geração anterior.

Isso significa que as *frequências genotípicas* alcançam uma relação de equilíbrio genético após uma geração de *cruzamentos aleatórios*, independentemente das *frequências genotípicas* da geração inicial, e se mantêm constantes ao longo do tempo, até que mudanças ocorram nas frequências dos alelos. Essa constância das *proporções genotípicas* em populações *panmíticas*, nas quais as *frequências gênicas* permanecem inalteradas, encerra o *conceito de Hardy-Weinberg*.

Em grandes populações *panmíticas*, as *frequências genotípicas* alcançam o *equilíbrio de Hardy-Weinberg* após uma geração de reprodução aleatória.

Ao rever a distribuição das *frequências gênicas* e *genotípicas* das 3 populações descritas na Tabela 8.1, verificamos que, apesar de elas apresentarem *frequências genotípicas* diferentes, todas alcançarão o *equilíbrio de Hardy-Weinberg* com proporções genotípicas de 0,36 L^ML^M: 0,48 L^ML^N: 0,16 L^NL^N, considerando que as *frequências alélicas* são iguais nos 3 d*emes*. Observe que a população III já se encontra em equilíbrio. Nas demais, ele será alcançado após uma geração de *panmixia* e, a partir daí, as *frequências genotípicas* permanecerão inalteradas se prevalecerem as condições iniciais, ou seja, o cumprimento dos pressupostos da *lei de Hardy-Weinberg*.

Pela lei de *Hardy-Weinberg*, em grandes populações diploides que se reproduzem sexual e panmiticamente, na ausência de *forças evolutivas*, a herança mendeliana mantém as *frequências gênicas* constantes, determinando as *proporções genotípicas*, geração após geração.

Testando o equilíbrio de Hardy-Weinberg | Teste do qui-quadrado (χ^2)

Vimos, portanto, que, de acordo com o *modelo de Hardy-Weinberg*, a partir das *frequências alélicas* em uma geração de uma população é possível prever as *frequências genotípicas* na geração seguinte e, estando a população em equilíbrio, as *frequências gênicas* e *genotípicas* permanecerão constantes. Mas como avaliar se as frequências dos genótipos observados em uma população se distribuem de acordo com o *equilíbrio de Hardy-Weinberg*? Retomando o exemplo apresentado, a geração inicial era constituída por 420 indivíduos L^ML^M, 600 L^ML^N e 180 L^NL^N. Sendo as frequências dos alelos L^M e L^N iguais a 0,60 (p) e 0,40 (q), esperaríamos que as frequências dos 3 genótipos (L^ML^M, L^ML^N e L^NL^N) no equilíbrio fossem p^2, $2pq$ e q^2, respectivamente, ou seja, 0,36, 0,48 e 0,16. Assim, na amostra total dos 1.200 indivíduos analisados, deveríamos ter 0,36 × 1.200 = 432 indivíduos com o genótipo L^ML^M, 0,48 × 1.200 = 576 com o genótipo L^ML^N e 0,16 × 1.200 = 192 com o genótipo L^NL^N. Ao comparar os dados dos genótipos observados na amostra com os esperados, verificamos que eles diferem. Logo, devemos ser capazes de avaliar a magnitude dessa diferença, ou seja, se o *desvio* observado é decorrente ou não do acaso. Ao supor que os dados obtidos se ajustam a um valor esperado, estabelecemos uma hipótese, denominada *hipótese nula* ou H_0, pois partimos da suposição de que não há diferença entre os valores observados e os preditos, de modo que quaisquer *desvios* encontrados poderão ser atribuídos ao acaso. Um teste estatístico que torna possível verificar se essas diferenças são significativas ou não é o teste do qui-quadrado (χ^2), que possibilita validar a *hipótese nula* para um conjunto de dados. Ele nos informa quanto os valores observados se ajustam aos esperados, indicando a probabilidade de que os *desvios* entre esses valores sejam decorrentes do acaso. Assim, podemos rejeitar ou não a *hipótese nula*. Se H_0 for aceita, significa que os *desvios* entre os dados observados e esperados são atribuídos ao acaso. Contrariamente, se H_0 for rejeitada, significa que os *desvios* não são decorrentes unicamente do acaso.

Qui-quadrado (χ^2)
Teste estatístico utilizado para checar o grau de concordância dos dados observados com as previsões de uma hipótese

O teste do χ^2 requer a utilização de dados numéricos, em vez de proporções, o que torna necessário fazer as devidas conversões. Seu valor corresponde a:

$$\chi^2 = \Sigma \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}} \qquad \text{(Equação 8.7)}$$

sendo, observado = número observado de indivíduos com determinado genótipo

esperado = número esperado de indivíduos com determinado genótipo

Σ = soma de todas as classes genotípicas

Conforme detalhado na Tabela 8.5, aplicando a Equação 8.7 ao nosso exemplo, obtivemos um valor de χ^2 de 2,083.

■ **Tabela 8.5** Aplicação do teste do qui-quadrado para testar se a distribuição de genótipos do grupo sanguíneo MN em uma amostra de 1.200 indivíduos norte-americanos está em *equilíbrio de Hardy-Weinberg*.

	Genótipos					
	L^ML^M		L^ML^N		L^NL^N	Total
Observado	420		600		180	1.200
Esperado	432		576		192	1.200
χ^2	$\frac{(420-432)^2}{432}$	+	$\frac{(600-576)^2}{576}$	+	$\frac{(180-192)^2}{192}$	
	0,333	+	1	+	0,75 = 2,083	

Grau de liberdade (*gl*)
Índice associado aos dados independentes que são utilizados na estimativa de um parâmetro. Em relação ao qui-quadrado para testar o equilíbrio de Hardy-Weinberg, *gl* corresponde ao número de classes dos dados menos o número de parâmetros estimados menos um

Após o cálculo do χ^2, é necessário saber interpretá-lo. Todo valor de χ^2 está associado a um número denominado grau de liberdade (*gl*), o qual, em relação ao *equilíbrio de Hardy-Weinberg*, corresponde ao número de categorias dos dados menos o número de parâmetros estimados a partir deles menos um.

Assim, tendo sido o parâmetro *p* estimado com base nos dados, *q* será automaticamente determinado, pois $q = 1 - p$, de maneira que apenas um desses parâmetros (*p*) foi diretamente estimado com base nos dados e, por isso, apenas um *grau de liberdade* relativo a ele será subtraído. Em nosso exemplo, temos 3 classes de dados possíveis (correspondentes aos 3 genótipos) e um parâmetro estimado (*p*), sendo o número de *graus de liberdade*, portanto, igual a $3 - 1 - 1 = 1$.

Após a obtenção do valor do χ^2 e do grau de liberdade a ele associado, o próximo passo é interpretar o χ^2 em termos de uma probabilidade (*P*) correspondente, que é obtida pela consulta à tabela padrão do qui-quadrado (Tabela 8.6). *P* representa a probabilidade de o acaso ser responsável pelo *desvio* entre os valores observados e esperados. Desse modo, um valor de χ^2 associado a uma probabilidade alta indica que o acaso é o principal responsável pelo *desvio* encontrado, o que aumenta nossa confiança na validade da hipótese testada. De maneira oposta, se a probabilidade associada ao χ^2 for baixa, significa que o acaso, por si só, dificilmente ocasionou o *desvio*, diminuindo a confiança na validade de nossa hipótese, levando a supor que algum outro agente, além do acaso, originou o *desvio*. Em geral, usa-se 0,05 (5%) como limite entre probabilidades altas e baixas, de maneira que quando $P \geq 0,05$, os dados mostram boa aderência ao modelo, não havendo razão para rejeitar a hipótese nula. Logo, há um indicativo de que o *desvio* observado no conjunto de dados ocorreu por influência do acaso. Se $P < 0,05$ significa que o *desvio* entre os valores observados e esperados é substancial, sendo ele decorrente do acaso em menos de 5% das vezes, indicando que devemos rejeitar a *hipótese nula*. Assim, quando $P < 0,05$, a probabilidade é estatisticamente significativa, sendo a hipótese inválida para os dados. Alternativamente, se $P \geq 0,05$, a hipótese é adequada, o que nos possibilita aceitá-la.

Tabela 8.6 Valores críticos da distribuição de χ^2, correspondentes a graus de liberdade (*gl*) específicos (primeira coluna à esquerda) e a diferentes probabilidades (*P*).

	Probabilidade (*P*)									
gl	0,99	0,90	0,70	0,50	0,20	0,10	0,05	0,01	0,002	0,001
1	0,0002	0,016	0,148	0,455	1,642	2,706	3,841	6,635	9,550	10,827
2	0,020	0,211	0,713	1,386	3,219	4,605	5,991	9,210	12,429	13,815
3	0,115	0,584	1,424	2,366	4,642	6,251	7,815	11,345	14,796	16,266
4	0,297	1,064	2,195	3,357	5,989	7,779	9,488	13,277	16,924	18,467
5	0,554	1,610	3,000	4,351	7,289	9,236	11,070	15,086	18,907	20,515
6	0,872	2,204	3,828	5,348	8,558	10,645	12,592	16,812	20,791	22,457
7	1,239	2,833	4,671	6,346	9,803	12,017	14,067	18,475	22,601	24,322
8	1,646	3,490	5,527	7,344	11,030	13,362	15,507	20,090	24,352	26,125
9	2,088	4,168	6,393	8,343	12,242	14,684	16,919	21,666	26,056	27,877
10	2,558	4,865	7,267	9,342	13,442	15,987	18,307	23,209	27,722	29,588
11	3,053	5,578	8,148	10,341	14,631	17,275	19,675	24,725	29,354	31,264
12	3,571	6,304	9,034	11,340	15,812	18,549	21,026	26,217	30,957	32,909
13	4,107	7,042	9,926	12,340	16,985	19,812	22,362	27,688	32,535	34,528
14	4,660	7,790	10,821	13,339	18,151	21,064	23,685	29,141	34,091	36,123
15	5,229	8,547	11,721	14,339	19,311	22,307	24,996	30,578	35,628	37,697
16	5,812	9,312	12,624	15,338	20,465	23,542	26,296	32,000	37,146	39,252
17	6,408	10,085	13,531	16,338	21,615	24,769	27,587	33,409	38,648	40,790
18	7,015	10,865	14,440	17,338	22,760	25,989	28,869	34,805	40,136	42,312
19	7,633	11,651	15,352	18,338	23,900	27,204	30,144	36,191	41,610	43,820
20	8,260	12,443	16,266	19,337	25,038	28,412	31,410	37,566	43,072	45,315

Retomando nosso exemplo, de posse dos valores do $\chi^2 = 2{,}083$ e de *gl* = 1 e pela consulta à Tabela 8.6, podemos obter a probabilidade associada e avaliar a adequação dos dados observados à hipótese testada. Observe que, no nível de tolerância estabelecido (0,05), o valor do χ^2 com um *grau de liberdade* é 3,841. Sendo nosso valor (2,083) menor do que o valor crítico do χ^2 e associado a uma probabilidade maior do que 0,05, compreendida entre 0,20 e 0,10, esse dado revela haver boa concordância entre os valores observados e esperados, indicando que os desvios notados podem ser atribuídos a flutuações aleatórias. Logo, podemos concluir que as *frequências genotípicas* para o grupo sanguíneo MN nessa população estão de acordo com as *proporções de Hardy-Weinberg*. De maneira oposta, se o resultado do χ^2 indicasse um *desvio* significativo entre as *distribuições genotípicas* observadas e esperadas, saberíamos que os genótipos na população em estudo não estariam em equilíbrio genético e várias poderiam ser as razões contribuindo para isso. *Desvios* significativos na *distribuição genotípica* de uma população são indicativos de que uma ou mais condições do *modelo de Hardy-Weinberg* encontram-se comprometidas. A ausência de *panmixia*, bem como a ocorrência de *mutação*, *seleção natural*, *migração* ou *deriva gênica* poderiam estar atuando e moldando novos padrões de variação genética.

Algumas aplicações do princípio de Hardy-Weinberg

Além das inúmeras implicações do *modelo de Hardy-Weinberg* em relação à estrutura genética das populações, em algumas situações, sua aplicação é de grande utilidade. Vimos que, em condições de equilíbrio, as frequências dos genótipos são determinadas pelas *frequências gênicas* e essa relação é mostrada na Figura 8.3.

Observe que, para um *locus* com 2 alelos, a frequência de heterozigotos em uma população em *equilíbrio de Hardy-Weinberg* nunca excede 0,5 e ocorre quando as *frequências alélicas* são iguais, ou seja, $p = q = 0{,}5$. Note que a proporção de heterozigotos, comparada a de homozigotos, aumenta gradativamente à medida que os alelos recessivos se tornam mais raros. É possível também verificar que, quando a frequência de um alelo é alta, a maioria dos organismos é homozigota para o alelo em questão.

Uma aplicação do *princípio de Hardy-Weinberg*, de grande valor prático, refere-se à determinação das *frequências alélicas* em dadas circunstâncias. Vimos que o cálculo das *fre-*

A frequência máxima de heterozigotos ocorre quando $p = q = 0{,}5$

Quando a frequência de um alelo é elevada, a maioria dos organismos é homozigota

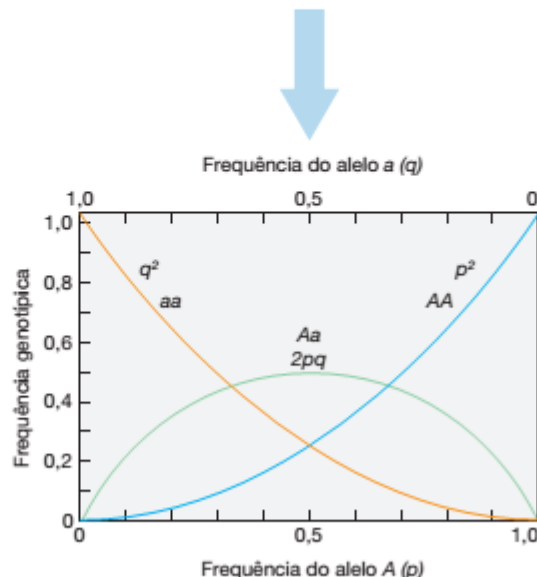



Figura 8.3 Relação entre as frequências alélicas e genotípicas no equilíbrio de Hardy-Weinberg para um *locus* com 2 alelos. As curvas representam as proporções dos genótipos homozigotos *A/A* (azul) e *a/a* (laranja) e heterozigotos *A/a* (verde) em populações com frequências gênicas diferentes, quando essas estão em equilíbrio.

quências gênicas com base nas *frequências genotípicas* é possível, desde que as frequências dos diferentes genótipos sejam conhecidas. Em relação ao traço usado como exemplo no decorrer deste capítulo (grupo sanguíneo MN), não há *dominância* entre os alelos, de maneira que os indivíduos heterozigotos podem ser diferenciados dos homozigotos (3 genótipos determinam 3 fenótipos). Entretanto, havendo *dominância*, isso não é possível, pois os heterozigotos não podem ser distinguidos dos homozigotos dominantes, de maneira que, heterozigotos e homozigotos dominantes constituirão uma classe fenotípica única (3 genótipos condicionam 2 fenótipos). Contudo, se os genótipos se distribuírem na população nas proporções do *equilíbrio de Hardy-Weinberg*, através da frequência dos homozigotos recessivos será possível estimar as *frequências alélicas*, aplicando-se a *equação de Hardy-Weinberg*, mesmo sem conhecermos as proporções dos demais genótipos na população. Sendo a frequência de homozigotos recessivos *aa*, no equilíbrio, equivalente a q^2, a frequência do alelo *a* (q) corresponderá à raiz quadrada desse valor, ou seja, à raiz quadrada da frequência dos genótipos homozigotos recessivos. Conhecendo-se o valor de q, será simples calcular a frequência do alelo dominante (p), pois $p = 1 - q$.

É bastante comum o interesse em conhecer a proporção de indivíduos heterozigotos para alelos recessivos em uma população, particularmente quando esses estão associados a doenças, considerando que os heterozigotos serão os transmissores desses traços. Nesse sentido, a aplicação da *lei de Hardy-Weinberg* possibilita uma boa estimativa dessa frequência, como veremos a seguir para a fibrose cística. A incidência desse distúrbio autossômico recessivo, entre europeus, é de 1/2.500 indivíduos, sendo a distinção entre homozigotos normais e heterozigotos possível somente a partir de análises moleculares do gene responsável pela doença, difíceis de serem conduzidas em larga escala. Estando a distribuição de genótipos neste *locus* em equilíbrio na população, a frequência do genótipo recessivo (*ff*), correspondente a q^2, servirá de base para o cálculo das frequências dos alelos,

$$q = \sqrt{f(aa)} \quad q = \sqrt{0{,}0004} = 0{,}02 \text{ e } p = 1 - q = 1 - 0{,}02 = 0{,}98$$

Conhecendo os valores de p e q, aplicamos a *equação de Hardy-Weinberg* (Equação 8.6) e estimamos as frequências de indivíduos homozigotos normais [$p^2 = (0{,}98)^2 = 0{,}9604$] e heterozigotos [$2pq = 2(0{,}98)(0{,}02) = 0{,}0392$] na população. Esse resultado evidencia que, aproximadamente, 4% dos indivíduos dessa população são heterozigotos e, portanto, transmissores da fibrose cística.

Outro exemplo de doença autossômica recessiva, mais rara que a fibrose cística, é a fenilcetonúria, decorrente da ausência da enzima fenilalanina hidroxilase, que prejudica o metabolismo do aminoácido fenilalanina e pode causar deficiência mental. Esse distúrbio acomete cerca de 1/10.000 indivíduos canadenses e, admitindo-se existência do *equilíbrio de Hardy-Weinberg*, podemos proceder ao cálculo das frequências gênicas.

Se $q^2 = 0{,}0001$, então $q = \sqrt{0{,}0001} = 0{,}01$

Logo, $p = 1 - q \rightarrow p = 1 - 0{,}01 = 0{,}99$

Com base no conhecimento das frequências dos alelos, aplicamos a Equação 8.6 e estimamos as frequências dos genótipos homozigotos dominantes [$p^2 = (0{,}99)^2 = 0{,}9801$] e heterozigotos [$2pq = 2\,(0{,}99)\,(0{,}01) = 0{,}0198$] na população.

Esses exemplos são oportunos para mostrar que: (1) alelos raros (aqueles cujas frequências na população são muito baixas), como os que causam a fenilcetonúria e a fibrose cística, comumente estão presentes em heterozigose, o que nos leva a inferir que, em geral, aqueles indivíduos que apresentam distúrbios recessivos são filhos de heterozigotos normais; (2) de maneira oposta, quando a frequência de um alelo é alta, ele ocorre em homozigose na maioria dos indivíduos da população, como mostrado na Figura 8.3.

O emprego da *lei de Hardy-Weinberg* possibilita calcular as *frequências gênicas* quando há dominância entre os alelos, desde que a população esteja em equilíbrio genético, sendo que a frequência do alelo recessivo será igual à raiz quadrada da frequência do fenótipo recessivo.

Quando um alelo recessivo é raro, a maioria dos indivíduos que portam esse alelo é de heterozigotos.

Extensões do equilíbrio de Hardy-Weinberg | Alelos múltiplos e genes ligados ao cromossomo X

Até aqui, vimos a aplicação da *lei de Hardy-Weinberg* para genes autossômicos com um par de alelos. No entanto, os princípios desse modelo de equilíbrio genético também são válidos para *loci* autossômicos com *alelos múltiplos* e para *genes ligados ao X*, desde que todos os pressupostos mencionados anteriormente sejam obedecidos. Vejamos como se distribuem as *frequências gênicas* e *genotípicas* no equilíbrio nessas situações.

Alelos autossômicos múltiplos

De acordo com o número de alelos presentes em um *locus*, novas variáveis podem ser introduzidas na equação que encerra o *equilíbrio de Hardy-Weinberg* (Equação 8.6), de maneira que esta passa a ser representada por $(p + q + r + \ldots + z)^2$. Se considerarmos a existência de um *locus* autossômico com 3 alelos (A_1, A_2 e A_3), nas respectivas frequências (p, q e r), sendo $p + q + r = 1$, as proporções no equilíbrio dos seis genótipos possíveis (A_1A_1, A_1A_2, A_1A_3, A_2A_2, A_2A_3 e A_3A_3) serão equivalentes a $(p + q + r)^2 = p^2 + 2pq + 2pr + q^2 + 2qr + r^2$ (Figura 8.4).

		Gametas masculinos		
	Alelos	A_1	A_2	A_3
	Frequências	p	q	r
Gametas femininos — Alelos	Frequências			
A_1	p	A_1A_1 p^2	A_1A_2 pq	A_1A_3 pr
A_2	q	A_2A_1 pq	A_2A_2 q^2	A_2A_3 qr
A_3	r	A_3A_1 rp	A_3A_2 rq	A_3A_3 r^2

Figura 8.4 Proporção de genótipos produzidos pela combinação de 3 alelos em um *locus* autossômico com base na união aleatória de gametas masculinos e femininos.

Em relação à existência de apenas um par de alelos autossômicos, vimos que populações cujas *frequências genotípicas* não estão em *equilíbrio de Hardy-Weinberg* alcançam o equilíbrio genético após uma geração de reprodução aleatória. Isso também se aplica aos *alelos autossômicos múltiplos*, ou seja, na ausência de *forças evolutivas* sistemáticas, as *frequências gênicas* permanecem constantes e as *frequências genotípicas* alcançam o equilíbrio após uma geração de *panmixia* e se mantêm inalteradas.

Consideremos o exemplo do grupo sanguíneo ABO, determinado pela combinação de 3 alelos autossômicos (I^A, I^B e i), sendo os alelos I^A e I^B *codominantes* entre si e dominantes em relação ao alelo i. Partindo de uma geração de uma grande população, que obedece aos pressupostos da *lei de Hardy-Weinberg* e na qual as frequências dos alelos I^A, I^B e i são iguais a $p = 0{,}3$, $q = 0{,}1$ e $r = 0{,}6$, respectivamente, após uma geração de *panmixia*, as *frequências genotípicas* se distribuirão de acordo com $(p + q + r)^2$ e alcançarão o equilíbrio, ou seja:

Grupo sanguíneo A → genótipos I^AI^A e I^Ai: $p^2 + 2pr = (0{,}3 \times 0{,}3) + 2 \times 0{,}3 \times 0{,}6 = 0{,}09 + 0{,}36 = 0{,}45$

Grupo sanguíneo B → genótipos I^BI^B e I^Bi: $q^2 + 2qr = (0{,}1 \times 0{,}1) + 2 \times 0{,}1 \times 0{,}6 = 0{,}01 + 0{,}12 = 0{,}13$

Grupo sanguíneo AB → genótipo I^AI^B: $2pq = 2 \times 0{,}3 \times 0{,}1 = 0{,}06$

Grupo sanguíneo O → genótipo ii: $r^2 = 0{,}6 \times 0{,}6 = 0{,}36$

Em relação a *alelos autossômicos múltiplos*, o *equilíbrio de Hardy-Weinberg* é alcançado após uma única geração de cruzamentos ao acaso.

Genes ligados ao X

As considerações relativas aos alelos autossômicos não podem ser completamente aplicadas às distribuições genotípicas de características ligadas ao cromossomo X, de maneira que, essas alcançam uma situação de estabilidade genética ligeiramente diferente dos traços autossômicos. De fato, somente no *sexo homogamético*, isto é, aquele que tem 2 cromossomos sexuais X, o cálculo das *frequências genotípicas* no *equilíbrio de Hardy-Weinberg* é feito da mesma maneira que o dos genes autossômicos. Sendo assim, para um *locus* do X com 2 alelos (X_A e X_a), as *proporções genotípicas* se distribuirão de acordo com $(p + q)^2 = p^2 + 2pq + q^2$ (Equação 8.6), sendo p^2, $2pq$ e q^2 as frequências dos genótipos X_AX_A, X_AX_a e X_aX_a, respectivamente, e p e q as frequências dos alelos na população total. Contudo, no *sexo heterogamético*, aquele *hemizigoto* para genes do X, as *frequências genotípicas* (X_AY e X_aY) no equilíbrio serão idênticas às *frequências gênicas* (p e q) da população (Figura 8.5).

		Gametas masculinos		
		Alelos no X		Alelos no Y
	Alelos	X_A	X_a	
	Frequências	p	q	
Gametas femininos – Alelos	Frequências			
X_A	p	X_AX_A p^2	X_AX_a pq	X_AY p
X_a	q	X_aX_A qp	X_aX_a q^2	X_aY q

Figura 8.5 Proporção de genótipos produzidos pela combinação de 2 alelos ligados ao cromossomo X com base na união aleatória de gametas masculinos e femininos.

Sendo assim, em relação aos alelos ligados ao X, a estabilidade da distribuição dos genótipos na população, como um todo, nem sempre é alcançada após uma geração de *cruzamentos aleatórios*. Isso ocorre somente quando as *frequências genotípicas* do *sexo heterogamético*, na geração inicial, correspondem às *frequências gênicas* da população total. Do contrário, várias gerações em *panmixia* são necessárias para que os genótipos se distribuam de maneira estável na população.

Suponhamos, em uma população humana, a ocorrência de um traço condicionado por um par de alelos, *A* e *a*, ligados ao X, com *frequências gênicas* p e q, respectivamente, lembrando que, em humanos, o sexo masculino é *hemizigoto* para alelos do cromossomo X. Os genótipos nessa população se distribuem nas seguintes frequências:

Mulheres $X_AX_A = 0{,}50$ $X_AX_a = 0{,}20$ $X_aX_a = 0{,}30$

Homens $X_AY = 0{,}60$ $X_aY = 0{,}40$

Aplicando-se a Equação 8.5 aos dados de nosso exemplo, obteremos as frequências dos alelos *A* e *a* na população total:

$$p = \frac{0{,}60 + 0{,}20 + 1{,}00}{3} = 0{,}60$$

$$q = \frac{0{,}40 + 0{,}20 + 0{,}60}{3} = 0{,}40$$

Ao verificarmos que as *frequências gênicas* na população ($p = 0{,}60$ e $q = 0{,}40$) são iguais às *frequências genotípicas* no sexo masculino, podemos antever que a distribuição estável dos genótipos segundo a *lei de Hardy-Weinberg* será alcançada após uma geração de *cruzamentos aleatórios*. Note que, em virtude das *frequências genotípicas* dos homens serem idênticas às *frequências gênicas* da população, uma única geração de *panmixia* será suficiente para

que as *frequências genotípicas* no sexo feminino passem a se distribuir como $p^2 + 2pq + q^2$ e permaneçam estáveis.

Se, em vez das *frequências genotípicas* citadas, a geração inicial da população fosse constituída por:

Mulheres $X_AX_A = 0{,}41$ $X_AX_a = 0{,}28$ $X_aX_a = 0{,}31$

Homens $X_AY = 0{,}70$ $X_aY = 0{,}30$

Teríamos $p = (0{,}70 + 0{,}28 + 0{,}82)/3 = 0{,}60$ e $q = 1 - 0{,}60 = 0{,}40$.

Isso nos possibilitaria concluir que a população não alcançaria o *equilíbrio de Hardy-Weinberg* após uma geração de *cruzamentos aleatórios*, pois as frequências dos alelos A e a (0,60 e 0,40, respectivamente), nesse caso, diferem das *frequências genotípicas* entre os homens (0,70 e 0,30). Desse modo, serão necessárias várias gerações de *cruzamentos aleatórios* para que o equilíbrio genético seja alcançado.

Traços recessivos ligados ao cromossomo X são mais comumente observados no *sexo heterogamético* (frequências q) do que no *homogamético* (frequência q^2). Como exemplos bem conhecidos de fenótipos dessa natureza na espécie humana, temos o daltonismo, a hemofilia A e a distrofia muscular de Duchenne. Sendo a frequência do alelo d (q) relacionada com o daltonismo, em determinada população, igual a 0,08 e estando ela em equilíbrio, essa também será a frequência de homens daltônicos. No entanto, a frequência de mulheres daltônicas nessa população será equivalente a $q^2 = (0{,}08)^2 = 0{,}0064$, ou seja, inferior a 1%. Em relação à hemofilia A, a frequência do alelo h (q) nesse mesmo *deme* é 0,0001, significando que 1/10.000 homens são hemofílicos. Entre as mulheres, por sua vez, a frequência da hemofilia A corresponde a $q^2 = (0,\ 0001)^2$, ou seja, essa doença é 1.000 vezes mais observada no sexo masculino do que no feminino.

A Tabela 8.7 resume a aplicação da *lei de Hardy-Weinberg* nas diferentes situações: um par de alelos autossômicos, alelos autossômicos múltiplos e alelos ligados ao cromossomo X.

■ **Tabela 8.7** *Lei de Hardy-Weinberg*: relação entre as frequências gênicas e genotípicas referentes a traços condicionados por 2 ou 3 alelos autossômicos e por genes ligados ao X.

Condição	Frequências gênicas	Frequências genotípicas
Dois alelos autossômicos	$f(A) = p$ $f(a) = q$	$f(AA) = p^2$ $f(Aa) = 2pq$ $f(aa) = q^2$
Três alelos autossômicos	$f(A_1) = p$ $f(A_2) = q$ $f(A_3) = r$	$f(A_1A_1) = p^2$ $f(A_1A_2) = 2pq$ $f(A_1A_3) = 2pr$ $f(A_2A_2) = q^2$ $f(A_2A_3) = 2qr$ $f(A_3A_3) = r^2$
Dois alelos ligados ao X	$f(X_A) = p$ $f(X_a) = q$	$f(X_AX_A) = p^2$ $f(X_AX_a) = 2pq$ $f(X_aX_a) = q^2$ $f(X_AY) = p$ $f(X_aY) = q$

Exceções ao modelo de Hardy-Weinberg

Veremos, a seguir, algumas situações, em que a *lei de Hardy-Weinberg* não é válida. Uma das condições fundamentais para a aplicação deste modelo é a *panmixia*, ou seja, a reprodução entre os membros da população deve ser ao acaso. Quando isso não acontece, a relação que o *princípio de Hardy-Weinberg* estabelece entre as *frequências gênicas* e *genotípicas* se desfaz.

É importante observar que, embora para muitos traços, incluindo o grupo sanguíneo MN, a escolha dos parceiros ocorra de maneira aleatória na população, ou seja, independentemente

de seu fenótipo, em algumas situações isso pode não necessariamente acontecer. Dois tipos de *cruzamento não aleatório* são conhecidos: a *reprodução preferencial*, pela qual os indivíduos de uma população selecionam seus parceiros com base em determinadas características fenotípicas, como altura, peso ou nível de escolaridade, e a *reprodução consanguínea* (*endocruzamento* ou *endogamia*), que se refere ao cruzamento entre organismos geneticamente aparentados. Nessas condições, em virtude de a reprodução não ocorrer de maneira randômica, as *frequências genotípicas* sofrem alterações com o passar das gerações e *desvios* em relação ao *equilíbrio de Hardy-Weinberg* são notados.

Uma consequência direta da *reprodução preferencial* e, particularmente, do *endocruzamento* na estrutura genética da população é o aumento da proporção de homozigotos em detrimento da proporção de heterozigotos, ou seja, populações endogâmicas apresentam frequências de genótipos homozigotos maiores em relação às populações *panmíticas*. O endocruzamento sistemático entre parentes próximos leva à homozigose total da população e constitui o tipo de *reprodução preferencial* que mais afeta as *frequências genotípicas*. Esse efeito, conhecido como *depressão por endocruzamento*, é nitidamente sentido por causar prejuízos, uma vez que aumenta a incidência de traços recessivos raros, muitos deles deletérios ou letais. A autofecundação, comum entre inúmeras espécies vegetais, constitui o tipo mais extremo de *endocruzamento*. Na espécie humana, uma consequência direta refere-se às taxas de mortalidade maiores e ao aumento de anomalias genéticas observadas na prole de uniões consanguíneas. Cabe salientar, entretanto, que em relação à criação e ao cultivo de diversas espécies de animais e vegetais, o *endocruzamento* é muitas vezes utilizado para obtenção de organismos com atributos pretendidos.

Endocruzamento (endogamia)
Tipo de reprodução preferencial que ocorre entre organismos geneticamente aparentados

Autofecundação
Fusão de gametas de um mesmo indivíduo

A maneira como os organismos escolhem seus parceiros, de maneira aleatória ou preferencial, é fator determinante na constituição dos genótipos de uma população, influenciando suas frequências.

É fácil verificar o efeito que a *endogamia* causa nas *frequências genotípicas*, pelas gerações, se conhecermos o coeficiente de endocruzamento (F), que expressa a intensidade do *endocruzamento* em uma população e cujo valor varia de 0 a 1. Como demonstrado na Tabela 8.8, os *desvios* nas *frequências genotípicas* decorrentes da *endogamia* podem ser medidos, sendo as alterações nas frequências diretamente proporcionais ao valor de F. Observe que ocorre redução na proporção de genótipos heterozigotos, em comparação às *frequências de Hardy-Weinberg*, e um aumento concomitante nas frequências de ambos os homozigotos à custa dos genótipos heterozigotos, de modo que, com o tempo, todos os indivíduos da população serão homozigotos. No caso extremo de uma população completamente *endogâmica*, ou seja, $F = 1$, podemos constatar que a frequência dos heterozigotos será equivalente a zero.

Coeficiente de endocruzamento
Probabilidade de um indivíduo ter 2 alelos iguais por herança

Tabela 8.8 Frequências genotípicas em populações que se reproduzem por endocruzamento.

Genótipo	Ocorrência de endocruzamento (F)	Frequências genotípicas: Cruzamento aleatório ($F = 0$)	Frequências genotípicas: Endocruzamento completo ($F = 1$)
AA	$f(AA) = p^2 + Fpq$	p^2	p
Aa	$f(Aa) = 2pq - 2Fpq$	$2pq$	0
aa	$f(aa) = q^2 + Fpq$	q^2	q

A *endogamia* utiliza a variação genética na população e a redistribui, favorecendo a homozigose, ou seja, modifica as frequências dos genótipos em relação às *proporções de Hardy-Weinberg*, conduzindo à perda da variabilidade genética.

Agentes promotores de variação genética nas populações

Em termos genéticos, vimos que os *demes* caracterizam-se pelas frequências de seus alelos em *loci* individuais em dado período de tempo. Entretanto, as populações não são unidades reprodutivas estáticas. Elas podem se expandir ou contrair, em virtude de taxas de natalidade e mortalidade, e trocar genes com outras populações pelo *fluxo migratório*. Assim, por seu caráter dinâmico, é comum que, ao longo das gerações, mudanças ocorram nos reservatórios

gênicos das populações e essas alterações, graduais e progressivas, sejam responsáveis por sua evolução.

A manutenção das *frequências gênicas* e *genotípicas*, observada sob condições de *equilíbrio de Hardy-Weinberg*, nos mostra que a reprodução sexuada, por si, não promove a evolução de um *deme*, uma vez que ela não é capaz de alterar as *frequências alélicas*. Observe que, embora a recombinação (*crossing over*) seja uma fonte importante de variabilidade genética, pois a cada geração origina novas combinações de alelos nos gametas, ela apenas rearranja a variação existente. Desse modo, o estado de equilíbrio genético, no que se refere à distribuição das *frequências gênicas* e *genotípicas* em uma população, estabelecido pelo *modelo de Hardy-Weinberg*, tende a ser mantido, geração após geração, até que ela sofra a ação de agentes capazes de modificar essas frequências. Mas, quais são os processos capazes de alterar o equilíbrio genético das populações? Forças como *mutação*, *migração*, *seleção natural* e *deriva genética aleatória*, denominadas *agentes evolutivos*, podem promover variações genéticas nas populações, necessárias ao processo evolutivo.

Recombinação (*crossing over*)
Formação de novas combinações gênicas, não encontradas nos genitores, como resultado da troca recíproca de material genético entre cromossomos homólogos durante a prófase I da meiose

Diferentemente do *cruzamento preferencial* – que, como vimos, pode alterar as *proporções genotípicas*, porém não altera as *frequências gênicas* –, e também do *crossing over* – capaz de rearranjar os alelos existentes embora não produza novos alelos –, as *forças evolutivas* são processos capazes de modificar, com diferentes intensidades, as *frequências alélicas* e *genotípicas* das populações. A seguir, veremos como cada um desses processos atua.

As populações evoluem por mudanças graduais em seus *reservatórios gênicos* e, para que isso ocorra, é necessária a ação de forças capazes de modificar as *frequências gênicas*.

Mutação

Mutação
Alteração ocorrida no DNA que constitui a principal fonte de variação genética

A mutação é a fonte original de variação genética, capaz de transformar um alelo em outro. Ela ocorre como um evento aleatório, que pode alterar, excluir ou duplicar qualquer segmento do DNA. Sendo as *mutações gênicas* alterações acidentais no material genético, é compreensível que, em sua maioria, elas sejam reconhecidas por seu efeito fenotípico deletério. Quando a *mutação* acontece na linhagem de células que dá origem aos gametas, ela pode ser transmitida aos descendentes e, dependendo da taxa com que ocorre, é possível que as *frequências alélicas* nos gametas não necessariamente reflitam aquelas da população original.

Para a maioria dos genes de animais e plantas, as *taxas de mutação* espontâneas são baixas, da ordem de 10^{-5} a 10^{-6} novas *mutações* por gameta por geração. Nesse sentido, *mutações* ocorridas em uma geração causam modificações muito pequenas nas frequências dos alelos na geração seguinte, constituindo, assim, uma *força evolutiva* fraca em termos de sua capacidade de alterar as *frequências gênicas*, uma vez que não promove mudanças imediatas na estrutura genética das populações.

Com base no conhecimento da *taxa de mutação* de um gene, é possível prever o impacto que ela, por si só, pode ocasionar nas *frequências alélicas* de uma geração para a próxima. Para ilustrar esse efeito, tomemos o exemplo de uma doença humana, autossômica dominante, condicionada pelo alelo *A*. Suponha a existência de 1.500.000 indivíduos em uma população específica, todos com genótipo *aa*, de maneira que a frequência inicial do alelo *a* é 1,0 e a frequência de *A* é zero. Essa geração inicial, portanto, é constituída por 3.000.000 gametas (2 × 1.500.000), todos portando o alelo *a*. Presumindo que a taxa de mutação (μ) de *a* para *A* é equivalente a $1{,}0 \times 10^{-6}$, o que significa que 1 em cada 1.000.000 alelos *a* mutará para o alelo *A*, o conjunto gênico dessa população, na geração seguinte, será constituído por 2.999.997 alelos *a* e 3 alelos *A* e, a cada geração, mantendo-se essa mesma taxa de mutação, 3 novos alelos *A* serão adicionados ao *pool gênico* em detrimento de 3 alelos *a* existentes (Tabela 8.9). Assim, partindo de uma frequência inicial do alelo *a* de 1,0 e de $\mu = 10^{-6}$, a frequência desse alelo diminuirá, gradualmente, e a do alelo *A* aumentará. Entretanto, com essa taxa de mutação serão necessárias milhares de gerações para que alterações apreciáveis nas frequências dos alelos *A* e *a* na população sejam notadas.

Fica evidente, portanto, que, em virtude das baixas taxas ocorridas na natureza, a *mutação* é capaz de alterar as proporções dos alelos em uma população de modo extremamente lento,

■ **Tabela 8.9** Introdução de novos alelos (*A*) por mutação direta ($\mu = 10^{-6}$) em uma população e taxas de mudança nas frequências gênicas observadas em gerações consecutivas.

Geração	Total de alelos *a* presentes na população antes da mutação	Alelos *A* surgidos por mutação	Total de alelos *a* após ocorrência de mutação	Frequência do alelo *a*	Frequência do alelo *A*
Inicial	3.000.000	0	3.000.000	1,0	0,0
1	3.000.000	3	2.999.997	0,999999	0,000001
2	2.999.997	3	2.999.994	0,999998	0,000002
3	2.999.994	3	2.999.991	0,999997	0,000003
4	2.999.991	3	2.999.988	0,999996	0,000004
5	2.999.988	3	2.999.985	0,999995	0,000005

Mutação direta
Mutação que transforma um alelo comum (selvagem) em um alelo raro (mutante), geralmente, deletério

Mutação reversa
O oposto da mutação direta. Um alelo mutante pode sofrer nova mutação e restabelecer a sequência gênica original, ou seja, o alelo selvagem

embora outros fatores também atuem nesse sentido. Não podemos desconsiderar o fato de toda modificação química envolvida na mudança de um alelo para outro ser recorrente, ou seja, poder se dar em ambos os sentidos ($A \leftrightarrow a$), embora, em geral, a mutação direta (μ) seja mais frequente que a mutação reversa (ν). Assim, mesmo ocorrendo mutações, se houver equivalência entre o número de *mutações diretas* e *reversas*, a mutação não irá alterar as *frequências gênicas*. Observe que, no exemplo dado (Tabela 8.9), não consideramos a ocorrência de *mutações reversas*, o que faria com que as mudanças nas proporções dos alelos fossem notadas ainda mais lentamente.

A *seleção natural* é outro processo que pode se contrapor à mutação, impedindo que, a partir dela, ocorram alterações nas *frequências gênicas*. É importante ter em mente que, de modo geral, novos alelos criados por *mutação* raramente aumentam o desempenho reprodutivo de um organismo ou conferem a ele melhor adaptação ao ambiente, de maneira que a manutenção desses alelos na população depende diretamente da *seleção natural*. Esta atua de maneira contínua contra os organismos portadores de genes deletérios que resultam em fenótipos mal adaptados ao ambiente.

Concluímos, portanto, que embora a *mutação* seja uma fonte genuína de variabilidade genética, por possibilitar que novos alelos sejam continuamente acrescidos ao *reservatório gênico* das populações, ela tem capacidade reduzida de alterar as *frequências gênicas*, de maneira imediata, se comparada aos outros *agentes evolutivos*.

Em decorrência das *taxas de mutações* espontâneas para a maioria dos genes serem baixas, elas promovem mudanças muito lentas nas frequências dos alelos nas populações.

Fluxo gênico

Fluxo gênico
Transferência de alelos de uma população para outra em virtude da migração de organismos férteis

Migração
Movimento de organismos de uma população para outra

Uma população pode variar de tamanho por causa de mudanças em suas taxas de natalidade ou de mortalidade e também em decorrência da incorporação ou subtração de organismos em virtude da *migração*. Pelo *fluxo migratório*, alelos podem ser acrescidos ao *conjunto gênico* de uma população, trazidos por organismos férteis oriundos de outros *demes*, ou perdidos, pelos organismos que emigram, de maneira que a *migração* constitui outro processo capaz de alterar as *frequências alélicas* e *genotípicas*. A introdução de alelos de uma população em outra, denominada fluxo gênico, é particularmente notada como importante fonte de variação genética se a população doadora, aquela da qual os migrantes se originam, e a receptora, que recebeu os migrantes, diferirem na sua composição genética, ou seja, apresentarem *frequências gênicas* diferentes. O efeito da migração sobre as frequências dos alelos depende de alguns fatores, especificamente: taxa de migração, ou seja, proporção de migrantes férteis que chegam à população receptora; tamanho da população receptora e magnitude da divergência genética entre as populações envolvidas, isto é, o quanto elas diferem em relação a suas *frequências alélicas*.

A Figura 8.6 ilustra o exemplo de duas populações humanas que habitam regiões geográficas distintas e que divergem em relação às frequências de 2 alelos de um gene autossômico, *A* e *a* ($p_1 \neq p_2$ e $q_1 \neq q_2$). Admita que, em determinada geração, tenha ocorrido a migração de um grupo

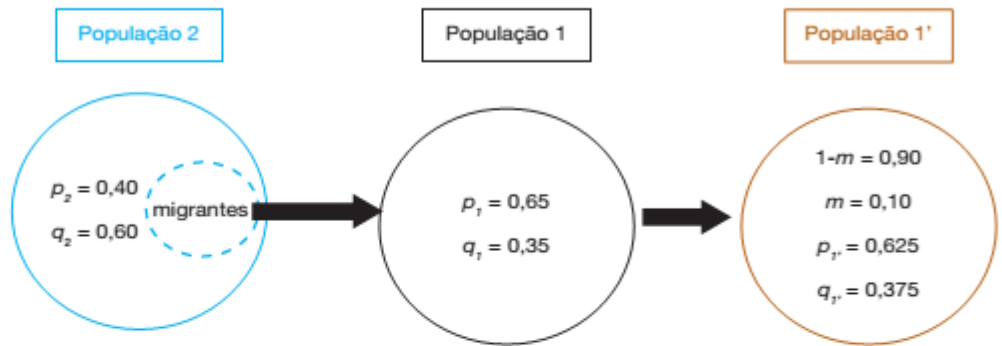



Figura 8.6 O esquema mostra a migração unidirecional de um grupo de indivíduos da população 2 para a população 1. O efeito do fluxo migratório entre demes depende da extensão da migração e das diferenças nas frequências alélicas.

de indivíduos da população 2 para a população 1, sem fluxo correspondente na direção contrária, ou seja, trata-se de uma *migração unidirecional*. Os gametas que eles carregam representam apenas uma amostra do *reservatório gênico* dos organismos reprodutores de sua população de origem. Suponha, ainda, que a proporção de migrantes oriundos da população 2 que chegam à população 1 corresponde a 10% do total de indivíduos da população mista (população 1′). Assim, essa população passa a ser formada por uma fração m de migrantes, sendo essa igual a 0,10, e por uma fração de indivíduos nativos, não migrantes, expressa como 1-m. Sob tais circunstâncias, as frequências dos alelos A e a na população 1′ são equivalentes a:

$$F(A) = p_{1'} = (1 - m)\,p_1 + mp_2$$

$$F(a) = q_{1'} = (1 - m)\,q_1 + mq_2$$

em que m = taxa de migração entre as populações

Considerando que as *frequências gênicas* são distintas nas duas populações originais de nosso exemplo, na geração seguinte, após o cruzamento aleatório entre o grupo de migrantes da população 2 com os nativos da população 1, o *deme* recém-constituído (população 1′) terá novas *frequências gênicas*. As proporções dos alelos A e a na população 1 (aquela que recebeu os imigrantes) correspondiam a $p_1 = 0{,}65$ e $q_1 = 0{,}35$, respectivamente, e as proporções desses alelos na população 2 (da qual os migrantes se originaram) eram iguais a $p_2 = 0{,}40$ e $q_2 = 0{,}60$, respectivamente. A população mista (população 1′), formada por 90% dos alelos da população 1 e 10% dos alelos da população 2, terá frequências alélicas, A ($p_{1'}$) e a ($q_{1'}$), equivalentes a:

$p_{1'} = (1 - 0{,}10) \times 0{,}65 + (0{,}10 \times 0{,}40) = 0{,}625$

$q_{1'} = (1 - 0{,}10) \times 0{,}35 + (0{,}10 \times 0{,}60) = 0{,}375$, o que também corresponde a 1- $p_{1'}$.

Observe que o fluxo gênico da população 2 para a população 1 mudou as frequências dos alelos A e a (p_1 e q_1) dessa população de 0,65 e 0,35 para 0,625 e 0,375, respectivamente, em apenas uma geração de *panmixia*. Isso mostra que quando a diferença entre as *frequências gênicas* das populações doadora e receptora for significativa, será prontamente observada mudança substancial nas frequências dos alelos.

A migração entre *demes* constitui fonte expressiva de variação genética.

Em geral, na natureza, *subpopulações* de uma espécie que habitam ambientes distintos tendem a diferir em relação a suas *frequências gênicas*, em virtude da adaptação ao meio. Os alelos trazidos pelos organismos que imigram, gradualmente, pela reprodução sexuada aleatória, são incorporados ao *reservatório gênico* da população receptora, de maneira que, a nova unidade reprodutiva resultante terá *frequências alélicas* intermediárias entre as frequências das populações originais (doadora e receptora). Desse modo, dependendo da intensidade com que ocorre, o *fluxo gênico* entre populações pode alterar as *frequências gênicas*, diminuindo a divergência genética entre as populações originais.

Se as *frequências alélicas* nas duas populações forem equivalentes, a *migração* não promoverá nenhum efeito na constituição genética da população recém-formada. É importante salientar também, que, nem sempre, os alelos transferidos pelos organismos que migram conferem a seus portadores uma boa adaptação ao novo ambiente, de maneira que, pela ação da *seleção*, eles podem ser naturalmente eliminados e não contribuir para o *pool gênico* da população.

Por constituir uma *força evolutiva* capaz de introduzir novos alelos em uma população, a *migração* assemelha-se à *mutação*. No entanto, diferentemente dessa, por ocorrer em uma taxa imensamente mais elevada em relação às taxas de *mutação gênica*, a *migração* é capaz de, prontamente, promover mudanças expressivas na estrutura genética das populações. Raros *demes* são sistemas fechados, nos quais não há *migração*. Em geral, a transferência gênica, em virtude do deslocamento de grupos de organismos entre *subpopulações* de uma espécie, mesmo entre aquelas geograficamente distantes, é uma possibilidade de ocorrência comum na natureza. Na história da humanidade, há relatos de inúmeros *fluxos migratórios* ocorridos que envolveram grupos de indivíduos que habitavam continentes diferentes.

A grande diferença entre a *mutação* e a *migração* refere-se à capacidade que a *migração*, em virtude das altas taxas de ocorrência, tem de alterar as frequências dos alelos. O efeito do *fluxo gênico*, como agente modificador, dependerá de quão divergentes forem as *frequências gênicas* das populações doadora e receptora.

Seleção natural

A imensa diversidade de genótipos existentes é uma forte evidência de que a variação genética é necessária ao *processo evolutivo*. Paralelamente à ação de agentes que são fonte de variabilidade na população, como a *mutação* e o *fluxo gênico*, a seleção natural atua, de maneira consistente, promovendo a *evolução adaptativa*. Ela é a principal força capaz de modificar as *frequências gênicas* em grandes populações.

Seleção natural
Força que opera, de maneira diferencial e não aleatória, selecionando os organismos de uma população com fenótipos mais bem adaptados a seu ambiente, ou seja, aqueles com maior êxito em relação à capacidade de sobrevivência e de reprodução

Sob o crivo da *seleção natural*, molda-se a variação genética observada na natureza, lembrando que a *seleção* atua sobre organismos, pois são eles que sobrevivem ou morrem e deixam maior ou menor descendência. Assim, ao discriminar fenótipos, a *seleção*, indiretamente, elege aqueles genótipos mais bem adaptados às condições ambientais existentes, o que leva a mudanças nas *frequências alélicas* pelas gerações de uma população.

Adaptabilidade e coeficiente de seleção

O ambiente no qual se vive limita o tamanho da população, levando os organismos a competirem por recursos que lhes possibilitem sobreviver. Nesse contexto, os organismos com características que proporcionam maior *sobrevida* e *desempenho reprodutivo* demonstram vantagem em relação a outros organismos de seu *deme*. Isso ocorre porque os genótipos desses organismos conferem a eles melhor adaptabilidade e, como resultado, seus alelos são preferencialmente transmitidos à geração seguinte e aumentam em frequência enquanto permanecerem as condições ambientais existentes. Assim, o conceito de *seleção natural* fundamenta-se no êxito diferencial de organismos com fenótipos mais adequados a seu ambiente imediato, os quais conseguem dar origem a um maior número de descendentes viáveis e férteis em relação àqueles menos adaptados.

Adaptabilidade
Componente hereditário do fenótipo que confere ao organismo vantagem em termos de sobrevivência e de reprodução em ambientes específicos

Como mostrado na Tabela 8.2, uma das condições da *lei de Hardy-Weinberg* diz respeito a todos os organismos da população serem igualmente viáveis e férteis. No entanto, na natureza, a *adaptação diferencial* dos organismos é incontestável. Alguns sobrevivem mais e mostram taxas de reprodução maiores em relação a outros, de maneira que essas vantagens individuais, em termos de *sobrevivência* e *fertilidade*, resultam em maior contribuição genética por parte desses organismos para as gerações futuras de sua população.

Os organismos mais adaptados são aqueles que expressam determinados *fenótipos vantajosos para a sobrevivência e a reprodução*, de modo que eles, em condições naturais de competição, serão selecionados em prejuízo dos demais. A adequação biológica que um fenótipo

Valor adaptativo (W)
Medida que expressa a contribuição de um organismo ao conjunto gênico da próxima geração (sucesso na transmissão gênica) em relação a outros organismos de sua população. Mede o sucesso reprodutivo relativo que determinado genótipo confere ao organismo

confere a um organismo em relação a outro, em um ambiente específico, é denominada valor adaptativo (W), sendo essa uma medida relativa que expressa a adaptação de um organismo às condições ambientais existentes em comparação aos demais organismos do *deme*. Assim, o *valor adaptativo* mede a capacidade que organismos com fenótipos distintos têm de transmitir seus alelos à geração seguinte.

Embora inúmeros sejam os fatores que influenciam a *adaptabilidade*, 2 componentes, *viabilidade* (*capacidade de sobrevivência*) e *sucesso reprodutivo*, são de extrema importância na adaptação de um fenótipo. Para que um organismo seja capaz de gerar prole numerosa é necessário que ele consiga sobreviver até a idade reprodutiva, seja fértil e alcance êxito reprodutivo. É claro que diferentes genótipos conferem aos organismos taxas diferenciadas de *desenvolvimento*, *sobrevivência* e *sucesso reprodutivo*, inclusive no que se refere à habilidade de atrair seus parceiros. Assim, cada fenótipo tem um *valor adaptativo* próprio que pode variar de 0 a 1, sendo, por definição, um traço de *valor adaptativo ótimo* todo aquele que aumenta as possibilidades de quem o tem, de contribuir com um maior número de descendentes para a geração seguinte. Desse modo, os seres capazes de produzir grande descendência são aqueles mais bem adaptados ao ambiente e, portanto, cujos genótipos têm *valor adaptativo* maior. Aqueles com genótipos que conferem baixo *sucesso reprodutivo* são menos adaptados, ou seja, têm *valor adaptativo* reduzido. A adaptação é função direta da capacidade que cada organismo tem de transmitir seus genes à próxima geração em relação aos demais organismos de seu *deme*.

Quando um genótipo confere *vantagem reprodutiva* a seu portador, ele tende a aumentar em frequência pelas gerações.

Coeficiente de seleção (*s*)
Medida que expressa a intensidade relativa da seleção natural contra diferentes fenótipos em virtude das diferenças adaptativas que eles conferem aos organismos da população

A intensidade de ação da *seleção natural* sobre fenótipos com diferentes graus de adaptação às condições existentes é medida pelo *coeficente de seleção* (*s*). Ele corresponde a $1 - W$ e expressa a redução na capacidade relativa de cada genótipo deixar descendentes em comparação com o genótipo mais adaptado (genótipo ótimo). Supondo que determinado traço genético tenha um *valor adaptativo* de 0,30, saberemos que o coeficiente de seleção relativo a esse traço será igual a 0,70, pois $s = 1 - 0{,}30$. Quando um fenótipo confere *valor adaptativo* nulo a seus portadores ($W = 0$), significando que eles não são capazes de se reproduzir e deixar descendentes (por serem estéreis ou não alcançarem a idade reprodutiva), isso significa que a *pressão de seleção* sobre esse fenótipo é total, ou seja, $s = 1$.

Em qualquer fase do ciclo de vida de um organismo diploide com reprodução sexuada, do zigoto à fase adulta, os efeitos da *sobrevivência diferencial* entre genótipos podem ocorrer e distorcer as *frequências genotípicas* na geração seguinte, ocasionando desvios no equilíbrio genético da população. Viabilidade desigual de zigotos, embriões, recém-natos e adultos, bem como taxas de fertilidade diferenciadas são comumente notadas entre genótipos distintos, indicando que a *seleção natural* pode atuar em diferentes níveis do desenvolvimento de um organismo. Além disso, o ciclo de vida dos organismos diploides inclui uma fase haploide, relativa à formação dos gametas, de maneira que a *seleção* pode também atuar nessa etapa (*seleção gamética*), considerando que os gametas nem sempre são igualmente viáveis. Assim, a *seleção* atua continuamente, preservando aqueles genótipos que resultam nos fenótipos mais adequados às condições ambientais existentes em determinado momento.

Novos alelos criados por *mutação* raramente melhoram a *adaptabilidade* do organismo, uma vez que, em geral, as condições ambientais existentes não são favoráveis aos fenótipos divergentes, de maneira que eles sofrem ação contínua da *seleção natural*. Em geral, alelos mutados propiciam efeitos fenotípicos deletérios que comprometem a adaptação do organismo a seu ambiente, prejudicando a continuidade da transmissão desse novo alelo à geração seguinte. Desse modo, os portadores de *mutações* deletérias, por mostrarem menor adequação biológica ao ambiente imediato, são selecionados em favor daqueles organismos sem tais mutações.

Penetrância completa
Refere-se aos alelos cuja expressão ao nível de fenótipo ocorre em 100% dos indivíduos portadores do gene

Embora a *seleção* atue no nível de fenótipo e não propriamente sobre os genes, no caso das *mutações* dominantes de penetrância completa, que expõem os organismos e genes à ação da *seleção* no momento em que surgem, os conceitos de *valor adaptativo* e de *coeficiente de seleção* podem ser estendidos aos alelos mutantes. Esse, porém, não é o caso dos genes cujos efeitos só são notados quando eles estão presentes em *homozigose*. Nessa situação, a seleção natural só atuará sobre os alelos recessivos expostos pelos fenótipos dos organismos *homozigotos*.

Nota-se existir um equilíbrio em relação ao número de alelos surgidos por *mutações* novas e aqueles perdidos por *seleção*.

É importante notar que os organismos estão em constante adaptação ao meio e que mudanças no ambiente podem modificar a *adaptabilidade* dos fenótipos, de modo que, um mesmo fenótipo pode ter *valores adaptativos* diferentes em ambientes distintos. Em determinadas circunstâncias, fenótipos que conferiam baixa *adaptabilidade* a seus portadores e sofriam forte pressão seletiva podem passar a ser positivamente selecionados.

O processo natural de *seleção* dos organismos com fenótipos mais bem adaptados às condições ambientais existentes promove, de maneira consistente, mudanças dirigidas nas *frequências alélicas* de uma geração para a seguinte.

Conhecendo-se as *frequências gênicas* e os *valores adaptativos* dos genótipos, é possível medir as alterações nas frequências dos alelos por causa da *seleção* em um *locus* específico, como veremos a seguir. Tomemos o exemplo de um traço condicionado por um par de alelos autossômicos (A_1 e A_2) com frequências 0,4 e 0,6, respectivamente (Tabela 8.10). O *valor adaptativo* de cada genótipo foi estimado dividindo-se o número médio de prole gerada pelo genótipo em questão pelo número de descendentes produzidos pelo genótipo mais fértil. Embora seja comum nos referirmos ao *valor adaptativo* relativo a um genótipo, cabe ressaltar, mais uma vez, que a *seleção natural* age sobre o organismo em sua totalidade.

Tabela 8.10 Mudanças notadas entre duas gerações nas frequências genotípicas de uma população em virtude da ação da seleção natural sobre genótipos com diferentes graus de adaptabilidade.

Genótipos	A_1A_1	A_1A_2	A_2A_2
Frequências genotípicas iniciais	$p^2 = (0,4)^2 = 0,16$	$2pq = 2\ (0,4)\ (0,6) = 0,48$	$q^2 = (0,6)^2 = 0,36$
Número médio de prole gerada	10	6	3
Valor adaptativo (W)	$W_{11} = 10/10 = 1,0$	$W_{12} = 6/10 = 0,6$	$W_{22} = 3/10 = 0,3$
Contribuição proporcional do genótipo para a população	$p^2W_{11} = (0,16)\ (1,0) = 0,16$	$2pqW_{12} = (0,48)\ (0,6) = 0,288$	$q^2W_{22} = (0,36)\ (0,3) = 0,108$
Frequências genotípicas relativas após a seleção	$\frac{p^2W_{11}}{\bar{W}} = 0,29.$	$\frac{2pqW_{12}}{\bar{W}} = 0,52$	$\frac{q^2W_{22}}{\bar{W}} = 0,19$

$$\bar{W} = p^2W_{11} + 2pqW_{12} + q^2W_{22} = 0,556$$

Frequências gênicas após a seleção: A_1 ($p = 0,55$) e A_2 ($q = 0,45$)

Observe que, de uma geração para a outra, em virtude da ação da *seleção natural* sobre os genótipos que conferem adaptação diferenciada ao ambiente, ocorreram mudanças nas *proporções genotípicas* e *gênicas*. As frequências dos 3 genótipos (0,16, 0,48 e 0,36), após a *seleção* ter atuado, mudaram para 0,29, 0,52 e 0,19, respectivamente. O alelo A_1, inicialmente em menor frequência (0,4), por conferir maior *adaptabilidade* a seus portadores em relação ao alelo A_2, na geração seguinte, teve sua frequência elevada para 0,55 em detrimento da frequência do alelo A_2, que diminuiu para 0,45.

A adaptação diferencial dos fenótipos às condições existentes, com o passar das gerações, ocasiona alterações nas *frequências gênicas* e *genotípicas* da população.

O exemplo da Tabela 8.10 mostra uma situação teórica, na qual alelos de um único *locus* influenciam diretamente a *adaptabilidade* do organismo ao seu ambiente, conferindo-lhes maior ou menor *sucesso reprodutivo* na transmissão de seus genes, como no caso de genes que conferem aos insetos resistência a inseticidas. Aqueles organismos com genótipos que aumentam a resistência, quando expostos a esse veneno, sobrevivem e perpetuam seus alelos em detrimento daqueles que não apresentam tal resistência, ou seja, os insetos resistentes manifestam vantagem seletiva na presença desse agente químico no ambiente.

Ainda em relação à preservação da variação genética em *loci* individuais, outras situações existem nas quais diferentes alelos são mantidos de maneira *balanceada*. Esse tipo de equilíbrio dinâmico se dá pela ação da seleção balanceadora, que favorece os heterozigotos em detrimento

Seleção balanceadora
Tipo de seleção natural que favorece os heterozigotos em um *locus*, em prejuízo dos homozigotos, em razão de eles apresentarem maior adaptação

de ambos os homozigotos. Isso ocorre em situações nas quais os indivíduos heterozigotos em um *locus* particular apresentam vantagens adaptativas em relação aos homozigotos, sendo essa superioridade conhecida como *vantagem do heterozigoto* ou sobredominância. Nesses casos, embora a *seleção* atue eliminando os alelos presentes em homozigose, pelos efeitos fenotípicos que causam, paralelamente, ela também age preservando esses mesmos alelos quando em heterozigose. Assim, 2 ou mais alelos de um *locus* podem ser mantidos na população, pela ação da *seleção balanceadora*, em proporções consideráveis, como *polimorfismo balanceado*.

Sobredominância
O heterozigoto demonstra maior adaptabilidade ao meio que ambos os homozigotos – a também chamada vantagem do heterozigoto

G·F Genética em Foco

Polimorfismo balanceado | O locus da β-globina

Um exemplo bem documentado de *polimorfismo balanceado* em humanos refere-se ao *locus* da cadeia da β-globina, um dos polipeptídios que formam a hemoglobina, proteína que transporta o oxigênio no sangue. Uma mutação no gene que codifica a β-globina (Hb^S) prejudica a formação das moléculas de hemoglobina a ponto de as hemácias assumirem a forma de foice. Indivíduos homozigotos para esse alelo (Hb^SHb^S) desenvolvem anemia falciforme, uma doença que pode levar à morte prematura, de modo que a *pressão de seleção* sobre eles, em relação aos homozigotos normais (Hb^AHb^A), é alta. Ocorre, no entanto, que em regiões onde a malária é endêmica, os indivíduos heterozigotos (Hb^SHb^A) são mais adaptados do que os Hb^AHb^A, uma vez que eles estão mais protegidos dos efeitos nocivos da malária, se infectados pelo agente transmissor *Plasmodium falciparum*, do que os indivíduos homozigotos normais (Hb^AHb^A). Logo, nessas regiões, a inferioridade de ambos os homozigotos (Hb^SHb^S e Hb^AHb^A) em relação aos heterozigotos (Hb^SHb^A), em termos de *adaptabilidade* às condições ambientais existentes, faz com que a seleção balanceadora atue mantendo ambos os alelos nesse *locus* como *polimorfismo balanceado*.

No entanto, a *adaptabilidade*, na maioria das vezes, não depende da ação de um único gene, mas da expressão de vários genes que, em conjunto, determinam os *traços quantitativos* responsáveis pela capacidade de *sobrevivência* e de *reprodução* do organismo. Assim, ao se tratar de *características quantitativas* e, considerando que fenótipos dessa natureza apresentam uma *distribuição contínua* (Figura 6.3), o efeito da *seleção natural* pode ser quantificado em termos do valor fenotípico médio da população. A maneira como o processo seletivo atua, favorecendo determinados fenótipos quantitativos em virtude das condições ambientais existentes, possibilita distinguir 3 tipos de seleção: *estabilizadora*, *direcional* ou *disruptiva* (Figura 8.7).

A *seleção estabilizadora* é aquela que favorece um conjunto de fenótipos bem adaptados às condições ambientais. Algumas populações, por existirem há várias gerações em um ambiente estável, podem alcançar um alto nível de adaptação, mantendo a variação genética do *deme* estável em um pico seletivo ou *norma adaptativa*. Assim, a seleção estabilizadora atua, continuamente, favorecendo *fenótipos intermediários*, associados à alta *adaptabilidade*, em detrimento daqueles fenótipos extremos que se desviam da norma e conferem baixa adaptação. Dados relacionando o peso de recém-natos com os índices de sobrevivência mostram ser o peso ótimo ao nascimento na espécie humana em torno de 3,5 kg. Bebês com pesos próximos a esse têm mais chances de sobrevivência em relação àqueles que se desviam, para mais ou para menos, desse valor.

Seleção estabilizadora
Processo seletivo que favorece fenótipos intermediários, bem adaptados ao ambiente, em prejuízo dos fenótipos extremos que conferem baixa adaptação

Em um ambiente estável, a maioria dos organismos está bem adaptada e novos tipos ou desviantes, provavelmente, terão adaptação menor.

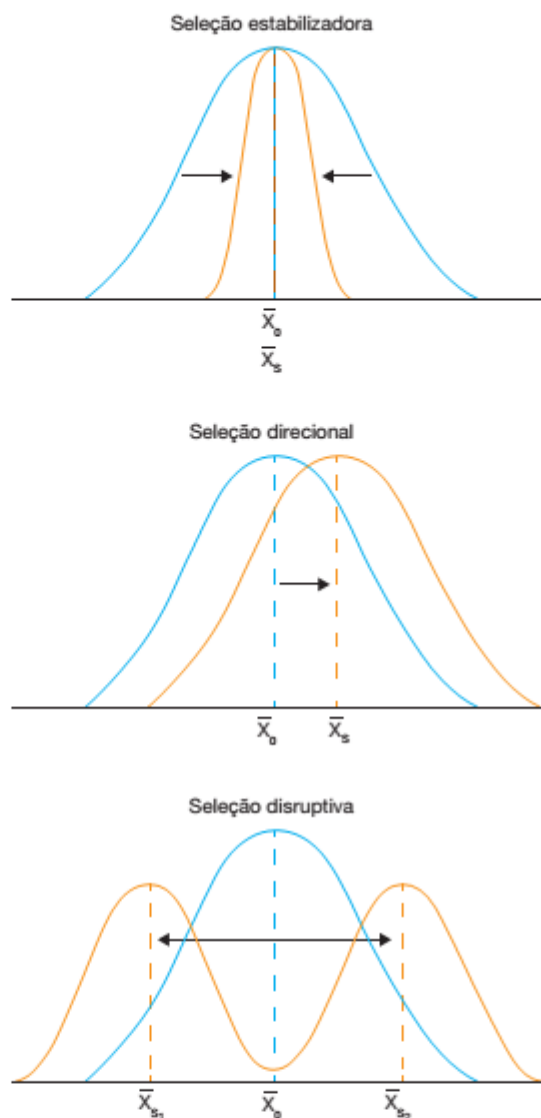



Figura 8.7 Efeito da seleção natural (estabilizadora, direcional e disruptiva) como agente modificador da distribuição de traços quantitativos na população. Nas 3 situações são apresentadas as médias populacionais (em azul) e as médias observadas após a seleção ter atuado (em laranja).

Seleção direcional
Força seletiva que favorece fenótipos em um extremo da distribuição

Seleção disruptiva
Processo seletivo que favorece fenótipos extremos, em prejuízo dos intermediários, em virtude de existir na população mais de um fenótipo bem adaptado às condições ambientais existentes

A **seleção direcional**, por sua vez, constitui um processo que ocorre quando o ambiente de uma população se modifica. Verifica-se, assim, uma mudança nos *valores adaptativos* dos fenótipos na população, em virtude das novas condições existentes, de maneira que, os organismos, até então favorecidos, passam a ser negativamente selecionados em favor daqueles localizados no extremo oposto da distribuição. Nota-se, portanto, mudança na ação da *seleção* em relação aos fenótipos mais bem adaptados, de uma direção para a outra da curva, de maneira que ela passa a favorecer aqueles organismos com características no outro extremo da distribuição fenotípica.

De maneira oposta à *seleção estabilizadora*, que elege os fenótipos intermediários, a **seleção disruptiva** favorece fenótipos bem distintos (*extremos fenotípicos*) em detrimento dos intermediários. Assim, quando as condições ambientais existentes favorecem a coexistência na população de mais de um traço quantitativo com *valor adaptativo* ótimo, a *seleção disruptiva* está atuando, elegendo os organismos situados nos 2 extremos da distribuição fenotípica, em prejuízo daqueles com fenótipos intermediários.

Apesar da necessidade de que toda população tem de estar adaptada a seu ambiente imediato, é igualmente importante que seu *reservatório gênico* mantenha suficiente variabilidade, de maneira a garantir sua não extinção diante das mudanças ambientais futuras.

Deriva genética

Embora a transmissão gênica de uma geração para a seguinte aconteça de maneira previsível, existe a possibilidade de que alguns eventos não previstos ocorram e interfiram nessa transmissão, ou seja, *agentes aleatórios* podem agir e promover mudanças na composição genética da população, processo ao qual denominamos deriva genética. A *deriva genética* é capaz de modificar as frequências dos alelos em uma população em relação à geração parental de modo indeterminado, ou seja, ao acaso, constituindo um mecanismo que reduz a variabilidade genética nas populações.

Deriva genética
Variação aleatória nas frequências gênicas em populações pequenas, causadas por erros de amostragem

Se considerarmos várias populações, todas de mesmo tamanho e com as mesmas *frequências gênicas*, sob efeito da *deriva genética*, com o passar das gerações ocorrerão *flutuações imprevisíveis* nas proporções de seus alelos, de modo que as populações se tornarão geneticamente distintas. Além disso, alelos poderão ser perdidos ou fixados de maneira aleatória, independentemente da adaptação que conferem aos organismos. É importante notar que a suscetibilidade de uma população à *deriva genética aleatória* está diretamente relacionada a seu tamanho. Para compreender melhor esse efeito, experimente arremessar uma moeda para o alto 10 vezes consecutivas. Embora o previsto seja obter cara em metade dos lançamentos e coroa na outra metade, é possível o resultado de 4 caras e 6 coroas ou mesmo 3 caras e 7 coroas, sem que isso cause grande surpresa. No entanto, se em vez de 10 vezes, você lançar a moeda 1.000 vezes, deverá obter uma proporção de caras e coroas muito próxima do esperado (0,50: 0,50). Caso o resultado seja 300 caras e 700 coroas, isso certamente será um forte indicativo de que algo está ocorrendo e fará você suspeitar da moeda. Torna-se claro, assim, que quanto menor o número de arremessos da moeda, maior o risco de que o acaso intervenha e promova um *desvio* no resultado previsto. Isso significa que o efeito da casualidade tem maior impacto nas pequenas amostras em relação às maiores. O *desvio* de uma proporção esperada em virtude do tamanho restrito da amostra é conhecido como erro de amostragem.

Erro de amostragem
Desvio da proporção esperada, decorrente do acaso, em virtude do número de eventos ser pequeno

Note que os gametas formados a cada geração transportam apenas uma amostra dos alelos presentes no *reservatório gênico* da população, ou seja, não representam integralmente os alelos existentes nos genitores. A cada geração temos um conjunto aleatório de $2n$ gametas extraídos da geração anterior de n indivíduos, os quais estão sujeitos a *erros de amostragem*. Assim, 2 alelos autossômicos, A e a, com frequências p e q, respectivamente, presentes em uma geração de uma população, têm 95% de chance de se manter na geração seguinte com frequências equivalentes a $p \pm 2\sigma$ e $q \pm 2\sigma$, ou seja, proporções que variam entre 2 desvios padrão (σ) acima e abaixo de suas respectivas frequências na geração anterior. Sendo o desvio padrão de p ou de q igual a $\sqrt{pq/2n}$, torna-se claro que os *desvios* observados nas frequências dos alelos A e a entre uma geração e outra dependem diretamente do tamanho da população (n).

A *deriva genética* ocorre quando o número de organismos férteis em uma população torna-se reduzido, não garantindo integralmente que os alelos do *pool gênico* sejam repassados à geração seguinte, nas proporções existentes. Em decorrência unicamente do acaso, alelos podem estar mais ou menos representados no *conjunto gamético*, ocasionando alterações aleatórias nas *frequências gênicas* da população. Portanto, em um *deme*, quanto menor o número de indivíduos capazes de se reproduzir, maiores serão os efeitos decorrentes do acaso nas *frequências alélicas*, de maneira que grandes desvios poderão ser notados em relação à composição genética, entre as gerações da população.

Embora nas populações infinitamente grandes o efeito da *deriva genética* seja mínimo, nas de pequeno tamanho ela pode constituir a principal *força evolutiva*, considerando que a variação aleatória nas *frequências gênicas*, promovida no decorrer das gerações, pode, casualmente, levar à fixação ou à perda de alelos. Portanto, nas *pequenas unidades reprodutivas*, verifica-se um decréscimo da variabilidade genética, com a fixação de alguns alelos e a eliminação de outros, em decorrência da *deriva genética aleatória*. Seria razoável supor que, nessas populações, a *mutação recorrente* pudesse repor os alelos perdidos pela *deriva genética*; no entanto, essa é uma possibilidade muito remota em razão das ínfimas taxas com que a *mutação* ocorre.

A *deriva genética* reduz a variação genética nas populações e quanto menor for o tamanho efetivo do *deme*, maior serão as flutuações nas *frequências alélicas*.

Todas as populações naturais estão sujeitas a *erros de amostragem* que podem ocasionar oscilações inesperadas nas frequências de seus genes de uma geração para a seguinte, sobretudo os *demes* pequenos. Assim, diferentes circunstâncias que acarretam a diminuição no tamanho populacional podem implicar *deriva genética* e, em algumas situações específicas, um efeito expressivo na estrutura genética da população será notado, como veremos a seguir.

Efeito fundador

Efeito fundador
Refere-se ao estabelecimento de uma população, com base em um pequeno grupo de organismos, cujos genótipos contêm apenas uma fração dos alelos da população original. Constitui causa de deriva genética

O efeito fundador é uma circunstância que causa *deriva genética* e ocorre quando alguns organismos de uma grande população emigram ou, por outro fator qualquer, separam-se de sua população de origem e fundam uma nova população. Assim, esse pequeno grupo de organismos, isolado de sua população nativa, passará a constituir um novo *deme*, se reproduzirá e crescerá em tamanho (Figura 8.8). O *conjunto gênico* agora existente é derivado, portanto, dos alelos presentes nos fundadores que, não necessariamente, são representativos do *pool* gênico da população da qual se originaram. É possível que, entre os fundadores, alguns organismos sejam mais férteis e contribuam com maior prole, de maneira que seus alelos serão transmitidos a muitos descendentes, o que levará ao aumento de suas frequências nesse novo *deme* ao longo das gerações.

Proporções alélicas observadas em certas populações humanas evidenciam a nítida ocorrência de *erros de amostragem*, ressaltando a ação da *deriva genética* como *força evolutiva*. Nessas populações, as frequências observadas são forte indicativo de que elas tenham se originado de um número reduzido de indivíduos fundadores, que constituíram pequenos isolados genéticos, separados por barreiras geográficas ou culturais. Um exemplo bem conhecido diz respeito à distribuição dos alelos que determinam o grupo sanguíneo ABO entre índios não miscigenados da região amazônica, os quais são isogênicos para o grupo O, ou seja, apenas o alelo i está segregando nessa comunidade (os alelos I^A e I^B inexistem). É bastante plausível, também, que o *efeito fundador* seja o provável responsável pela incidência relativamente elevada de algumas doenças hereditárias raras em populações humanas contemporâneas, como o albinismo entre os habitantes da ensolarada Ilha dos Lençóis, a 200 km da costa do Maranhão; a doença de Tay-Sachs entre judeus Ashkenazi e a distrofia diastrófica entre finlandeses.

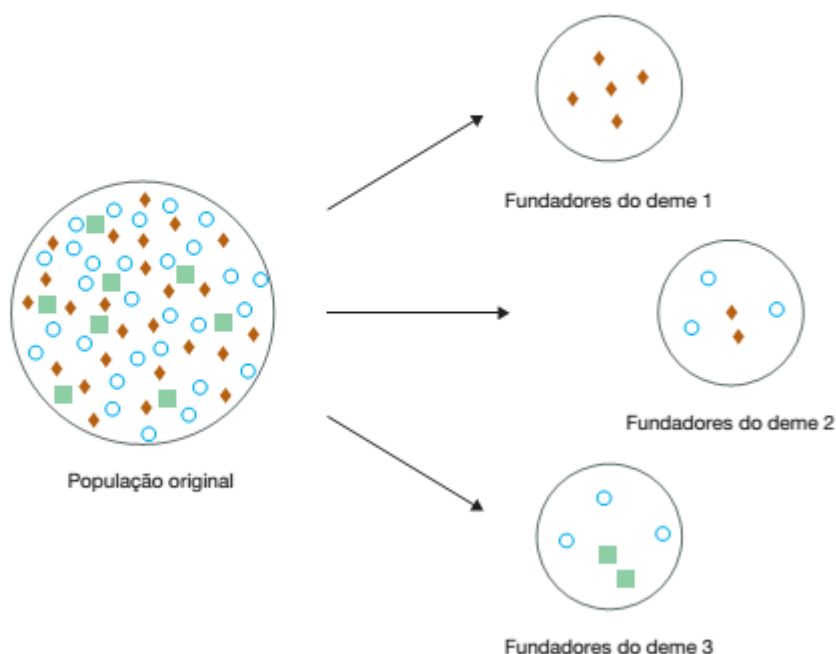



Figura 8.8 Variações são notadas na composição genética de populações quando elas se originam de um pequeno grupo de organismos. Nas 3 situações mostradas observa-se que os genótipos fundadores dos demes 1, 2 e 3 não representam integralmente o *pool* gênico da população da qual se originaram.

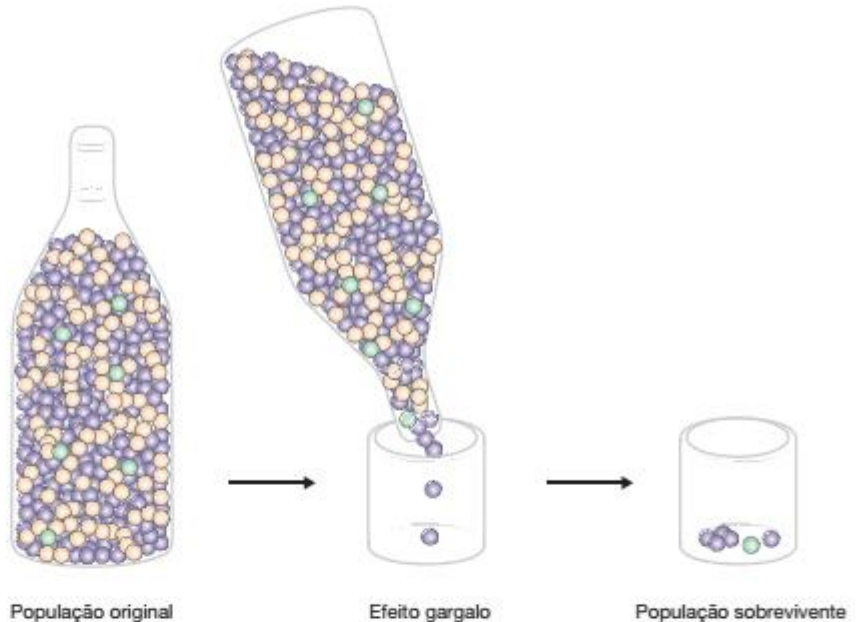
Gargalo genético

Gargalo genético
Redução drástica no tamanho da população, com consequente perda da variabilidade genética, que leva à deriva genética

Outro fator que pode levar à *deriva* é o gargalo genético, que ocorre quando o tamanho de uma grande população é drasticamente reduzido. A diminuição acentuada de uma população pode se dar em virtude de acontecimentos repentinos, como grandes epidemias e outras catástrofes naturais (enchentes), capazes de possibilitar a sobrevivência de um pequeno número de indivíduos reprodutivos. Esses desastres, ainda que transitórios, podem reduzir seriamente a *diversidade genética* da população.

O *efeito gargalo* decorre do estreitamento na passagem de genes de uma geração para outra em uma população que, em razão da diminuição abrupta e aleatória de seu tamanho, passou por drástica redução de genitores. Assim, a geração seguinte, originada a partir do conjunto de gametas dos organismos sobreviventes, poderá apresentar composição genética distinta da população pré-gargalo, que não necessariamente corresponde às *frequências alélicas* do *pool gênico* original. Em decorrência disso, haverá decréscimo na variabilidade genética populacional, como mostrado na Figura 8.9. No *deme pós-gargalo*, pelo efeito do acaso, alguns alelos podem se disseminar e aumentar suas frequências, ao passo que outros podem estar ausentes ou em menor frequência.

A *deriva genética* promove oscilações aleatórias nas *frequências gênicas*, ao longo das gerações, alterando a composição genética da população. Nota-se redução da variação genética em virtude da fixação ou perda de alelos.

Figura 8.9 O que ocorre em uma população quando esta passa por uma redução drástica de tamanho em virtude de acontecimentos repentinos. As bolinhas de diferentes cores correspondem aos genótipos dos organismos. Observe que, na população sobrevivente, aleatoriamente, alguns genótipos estão super-representados e outros estão ausentes.

RESUMO

- O processo evolutivo envolve mudanças genéticas na estrutura das populações ao longo do tempo, de modo que, com o estudo da dinâmica dos genes nas populações e das forças que moldam a variação genética, é possível compreendê-lo
- Os organismos que formam unidades populacionais locais (subpopulações ou demes) compartilham um conjunto gênico e entrecruzam
- Denomina-se frequência alélica ou gênica à proporção dos alelos em um *locus* em uma população e a frequência genotípica à proporção dos genótipos resultantes da combinação dos alelos
- A lei de Hardy-Weinberg aplica-se a grandes populações de organismos diploides, com reprodução sexuada e aleatória, que não sofram a influência de forças evolutivas (mutação, migração, seleção natural e deriva genética)
- O princípio de Hardy-Weinberg estabelece relação de equilíbrio entre as frequências gênicas e genotípicas em grandes populações panmíticas: as frequências dos genótipos são previstas pelas frequências alélicas e vice-versa; na ausência de forças evolutivas sistemáticas, as frequências não se alteram de uma geração para a seguinte; e as frequências genotípicas alcançam o equilíbrio de Hardy-Weinberg

após uma geração de cruzamentos aleatórios e se mantêm constantes

- A constância na proporção dos alelos e dos genótipos na população implica que, na ausência de forças evolutivas capazes de modificar as frequências gênicas, o mecanismo da herança mendeliana, por si só, é capaz de manter o equilíbrio genético geração após geração
- O modelo de Hardy-Weinberg estabelece que quando as frequências genotípicas de uma população se apresentam de acordo com $(p + q + r + \ldots + z)^2$ ela está em equilíbrio genético, sendo este alcançado em apenas uma geração de panmixia. Para genes ligados ao X, a estabilidade na distribuição dos genótipos na população ocorrerá após uma geração de cruzamentos aleatórios quando as frequências genotípicas do sexo heterogamético forem iguais às frequências gênicas da população total. Do contrário, várias gerações serão necessárias
- Para um *locus* com 2 alelos, a frequência máxima de heterozigotos em uma população em equilíbrio corresponde a 0,5 e ocorre quando as frequências gênicas são iguais ($p = q = 0{,}5$)
- Quando a frequência de um alelo é alta na população, a maioria dos organismos é homozigota para esse alelo. Quando o alelo é raro, a maioria dos genótipos com o alelo é heterozigota
- Existem duas formas de cruzamento não aleatório: a reprodução preferencial e a reprodução consanguínea (endocruzamento ou endogamia). Ambas ocasionam aumentos nas proporções dos homozigotos em detrimento dos heterozigotos e levam à perda de variabilidade genética
- Um dos efeitos prejudiciais do endocruzamento refere-se ao aumento dos genótipos homozigotos para alelos raros e deletérios
- As forças capazes de alterar o equilíbrio e promover variações genéticas nas populações são mutação, migração, seleção natural e deriva genética
- A mutação é uma fonte genuína de variação genética. Entretanto, em virtude das baixas taxas de ocorrência, ela não promove mudanças imediatas nas frequências dos alelos em uma população
- A seleção natural é a principal força capaz de promover modificações dirigidas nas frequências gênicas pelas gerações de uma população. Ela atua continuamente, em qualquer fase da vida de um organismo, preservando aqueles mais bem adaptados ao ambiente
- A adaptação diferencial dos organismos às condições existentes é medida pelo *valor adaptativo* (W), que expressa a capacidade que um organismo demonstra de transmitir seus genes à geração seguinte em relação aos demais organismos da população. Dois componentes influenciam diretamente a adaptabilidade, a capacidade de sobrevivência e o sucesso reprodutivo
- Diferenças na sobrevivência e no sucesso reprodutivo de indivíduos com diferentes genótipos podem modificar as frequências alélicas
- O coeficiente de seleção (s) mede a ação da seleção natural sobre fenótipos com diferentes valores adaptativos
- A seleção balanceadora favorece os heterozigotos, quando esses apresentam vantagens adaptativas (sobredominância) em relação a ambos os homozigotos. Possibilita existência na população de polimorfismos balanceados
- A seleção estabilizadora favorece fenótipos quantitativos intermediários, bem adaptados às condições ambientais, em detrimento dos fenótipos extremos, que conferem baixa adaptação
- A seleção direcional atua quando mudanças drásticas ocorrem no ambiente e os valores adaptativos dos fenótipos mudam, de maneira que aqueles localizados no extremo oposto da distribuição fenotípica passam a ser favorecidos
- A seleção disruptiva atua quando as condições ambientais existentes favorecem a coexistência na população de mais de um traço quantitativo com valor adaptativo ótimo (extremos fenotípicos) em detrimento dos fenótipos intermediários
- O fluxo gênico é capaz de alterar as frequências alélicas e genotípicas pela exclusão ou inclusão de genes em uma população. O efeito da *migração* sobre as frequências dos alelos depende da taxa de migração, do tamanho da população receptora e da divergência genética entre as populações
- Uma importante diferença entre a mutação e a migração refere-se à capacidade que a migração, em virtude das altas taxas de ocorrência, tem de alterar as frequências gênicas
- A deriva genética, por ocorrer de maneira aleatória, causa oscilações substanciais nas frequências alélicas em populações pequenas, ao longo do tempo, podendo levar à perda ou à fixação de alelos. Quanto menor o tamanho da população, maiores os efeitos da deriva genética
- O efeito fundador e o gargalo genético são circunstâncias que podem levar à deriva genética e reduzir a diversidade genética da população
- O efeito fundador refere-se ao estabelecimento de uma nova população por um pequeno grupo de organismos fundadores
- O gargalo genético decorre do estreitamento na passagem de genes de uma geração para outra em uma população que sofreu diminuição drástica e aleatória de tamanho, com perda de variabilidade genética.

AUTOAVALIAÇÃO

8.1 Calcule as frequências gênicas e genotípicas no equilíbrio em relação a duas populações cujas frequências dos genótipos, determinados por um par de alelos autossômicos, são: $AA = 0{,}05$, $Aa = 0{,}30$ e $aa = 0{,}65$ (população 1) e $AA = 0{,}20$, $Aa = 0$ e $aa = 0{,}80$ (população 2).

8.2 A fenilcetonúria é uma grave doença metabólica de herança autossômica recessiva que pode ser detectada em bebês, por meio do teste do pezinho, alguns dias após o nascimento. Esse teste, realizado em 111.430 recém-natos de uma mesma população, identificou 10 casos com a doença. A partir desses dados, determine as frequências gênicas e genotípicas na população.

8.3 Quais são as frequências genotípicas no equilíbrio em uma população panmítica na qual a frequência do fenótipo autossômico recessivo é 8%?

8.4 O albinismo e a anemia falciforme são condições autossômicas recessivas. Em uma grande população panmítica, a frequência de pessoas albinas é de 1 em 20.000 e de indivíduos com anemia falciforme é igual a 1 em 625. Que percentual de indivíduos nessa população estima-se que sejam hetero-

zigotos para o gene do albinismo? E para o gene da anemia falciforme?

8.5 Duas populações mostram as seguintes frequências genotípicas:

População I	0,24 *AA*	0,30 *Aa*	0,46 *aa*
População II	0,33 *AA*	0,12 *Aa*	0,55 *aa*

Se os cruzamentos ocorrerem de maneira aleatória, quais serão as frequências dos 3 genótipos nas duas populações na próxima geração?

8.6 Considere uma amostra de indivíduos testada para o grupo sanguíneo MN na qual foram registrados 3.574 com o fenótipo MM, 6.078 MN e 2.606 NN. Calcule as frequências genotípicas e dos alelos M e N nessa amostra. Ela encontra-se em equilíbrio de Hardy-Weinberg?

8.7 Verifique se as populações II e III mostradas na Tabela 8.1 encontram-se em equilíbrio genético.

8.8 Entre 2.000 indivíduos testados para o grupo sanguíneo ABO, foram identificados 900 A, 260 B, 120 AB e 720 O. Calcule as frequências dos alelos I^A, I^B, i e dos genótipos possíveis desse sistema sanguíneo.

8.9 Em uma população, 2.000 indivíduos foram testados para o grupo sanguíneo ABO e foram identificados: 640 do grupo A, 300 do grupo B, 80 do grupo AB e 980 do grupo O. Determine a frequência dos alelos I^A, I^B, i e dos diferentes genótipos: I^AI^A, I^Ai, I^BI^B, I^Bi, I^AI^B e ii.

8.10 Um certo traço recessivo ligado ao X se distribui em uma população da seguinte maneira:

Mulheres:	$X^AX^A = 554$	$X^AX^a = 108$	$X^aX^a = 14$
Homens:	$X^AY = 622$	$X^aY = 84$	

Determine as frequências gênicas nessa população.

8.11 Avalie se as populações descritas a seguir, ambas com 1.000 indivíduos de cada sexo, encontram-se em equilíbrio de Hardy-Weinberg.

População I: $X^AY = 0{,}60$ $X^aY = 0{,}40$ $X^AX^A = 0{,}51$ $X^AX^a = 0{,}18$ $X^aX^a = 0{,}31$

População II: $X^AY = 0{,}70$ $X^aY = 0{,}30$ $X^AX^A = 0{,}43$ $X^AX^a = 0{,}24$ $X^aX^a = 0{,}33$

8.12 Dois em cada 100 homens, em uma determinada população em equilíbrio de Hardy-Weinberg, apresentam daltonismo, um distúrbio recessivo ligado ao X. Partindo dessa informação, determine a frequência de mulheres daltônicas e de heterozigotas na população.

8.13 A hemofilia é uma doença causada por um gene recessivo ligado ao X. A frequência desse alelo em uma grande população panmítica é de 0,02. Qual é a frequência esperada de homens e de mulheres hemofílicas na população?

8.14 Sendo a frequência de mulheres com um traço recessivo ligado ao X de 19,36%, e estando a população em equilíbrio, qual é a proporção de homens com o traço nesse deme?

8.15 Quais são as forças evolutivas capazes de modificar as frequências gênicas e genotípicas em uma população?

8.16 Comente: "A maioria das mutações novas é deletéria."

8.17 Defina valor adaptativo e cite quais são os principais componentes da adaptabilidade.

8.18 Sendo os valores adaptativos de 3 genótipos em uma população equivalentes a $W_{AA} = 0{,}9$, $W_{Aa} = 1{,}0$ e $W_{aa} = 0{,}7$ e a frequência do alelo A igual a 0,5, qual será a frequência desse alelo na geração seguinte?

8.19 Qual é o principal efeito da endogamia em uma população?

8.20 Caracterize e exemplifique seleção balanceadora.

8.21 O que é sobredominância?

8.22 Em relação aos valores adaptativos dos genótipos a seguir listados em 3 diferentes demes, interprete como a seleção natural está atuando (s e s_1 representam diferentes coeficientes de seleção, sendo $0 < s_1 < s < 1$). Em qual situação se constata a existência de sobredominância?

	AA	*Aa*	*aa*
Deme 1	1	$1 - s_1$	$1 - s$
Deme 2	$1 - s$	1	$1 - s_1$
Deme 3	1	1	$1 - s$

8.23 Diferencie seleção estabilizadora de seleção disruptiva.

8.24 Conceitue deriva genética. Ela pode levar à fixação em uma população de alelos que confiram a seus portadores baixo valor adaptativo? Por quê?

8.25 Comente o princípio fundador.

8.26 O que você entende por gargalo genético?

BIBLIOGRAFIA

BEIGUEMAN, B. *Dinâmica dos genes nas famílias e nas populações.* 1ª ed., São Paulo: Sociedade Brasileira de Genética, 1994.

CAMPBELL, N.A; REECE, J.B.; URRY, L.A. *et al. Biologia.* 8ª ed., Porto Alegre: Artmed, 2010.

FLETCHER, H.; HICKEY, I.; WINTER, P. *Genetics.* 3ª ed., Reino Unido: Taylor & Francis Group, 2007.

GRIFFITHS, A.J.F.; WESSLER, S.R.; LEWONTIN, R.C. *et al. Introdução à genética.* 9ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

HARTL, D.L.; CLARK, A.G. *Princípios de genética de populações.* 4ª ed., Porto Alegre: Artmed, 2010.

KLUG, W.S.; CUMMINGS, M.R.; SPENCER, C.A. *et al. Conceitos de genética.* 9ª ed., Porto Alegre: Artmed, 2010.

METTLER, L.E; GREGG, T.G. *Genética de populações e evolução.* 1ª ed., São Paulo: Polígono, 1973

PASTERNAK, J.J. *Uma introdução à genética molecular humana.* 2ª ed., Rio de Janeiro: Guanabara Koogan, 2007.

PIERCE, B.A. *Genética, um enfoque conceitual.* 3ª ed., Rio de Janeiro: Guanabara Koogan, 2011.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

9

Genética do Câncer

Objetivos de estudo

Compreender os mecanismos gerais de desenvolvimento do câncer
Definir e avaliar a importância dos diferentes tipos de mutação gênica
Entender as diferenças entre mecanismos genéticos e epigenéticos
Avaliar a importância dos mecanismos de controle da proliferação celular
Diferenciar agente mutagênico de agente carcinogênico, com exemplos
Definir oncogenes e genes supressores tumorais, com exemplos
Conhecer as principais síndromes do câncer hereditário
Perceber a importância da análise genética na prevenção e no tratamento do câncer

Conceitos-chave do capítulo

Adutos de DNA
Alelo
Amplificação gênica
Aneuploidia
Angiogênese
Apoptose
Camundongos *knockout*
Câncer
Células-tronco
Citocinas
Cromossomo Filadélfia
Deleção gênica
DNA polimerase
Enzimas
Fator de transcrição
Fatores de crescimento
Fenótipo
Herança poligênica
Homeostase
Inibição por contato
Iniciação
Matriz extracelular
Metástase
Mecanismos epigenéticos
Mitógeno
Mitose
Metilação do DNA
Mutação
Mutação de ponto
Mutação esporádica
Mutação germinativa
Mutação hereditária
Não disjunção
Oncogene
Progressão
Promoção
Proto-oncogene
Quinase
Receptores celulares
Retinoblastoma
Sarcoma
Senescência
Translocação gênica

Introdução

Câncer
Conjunto de doenças caracterizadas por crescimento celular descontrolado, morfologia do tecido alterada e capacidade de migração das células para outros órgãos

Homeostase
Manutenção das condições constantes do meio interno de um organismo

A principal característica do câncer é o crescimento celular descontrolado, ocasionando a perda do fenótipo da célula normal e formando um tecido diferente deste que a originou. Dentro de um contexto mais amplo, esse tecido age independentemente dos controles da homeostase para o crescimento normal, alterando sobremaneira a fisiologia do órgão, podendo resultar na morte do organismo portador. Na verdade, o câncer é um grupo heterogêneo de doenças. Ele é um dos maiores problemas de saúde pública no mundo, constituindo a segunda causa de morte em países desenvolvidos, sendo apenas superada pelas doenças cardiovasculares. No Brasil não é diferente, e a perspectiva é a de que a cada ano um grande número de indivíduos venha a óbito por essa doença tão assustadora. Mas o que dá origem à célula cancerosa? Quais são os mecanismos genéticos envolvidos em seu desenvolvimento? Como essas células são capazes de se estabelecer e crescer em outros órgãos? Essas e outras perguntas são a base deste capítulo que descreve os principais elementos da Genética envolvidos na formação do câncer.

■ O câncer como doença genética

Aneuploidia
Alterações no número normal de cromossomos em uma célula

Não disjunção
Separação inadequada dos cromossomos durante a meiose ou mitose

Mitose
Processo no qual ocorre a divisão da célula eucariótica, originando uma nova célula semelhante

Mutação
Alteração na sequência de nucleotídios do DNA

Mecanismos epigenéticos
Conjunto de processos moleculares que modificam a expressão dos genes sem alterar sua sequência de nucleotídios

Metilação do DNA
Processo de adição de grupamentos metila no DNA por meio de enzimas DNA metiltransferases

Há muito tempo os pesquisadores da área de câncer identificaram esta como uma doença ligada a alterações do material genético. Uma das observações mais frequentes é a forma diferente do núcleo e o número anormal dos cromossomos das células de um tumor maligno. A aneuploidia, portanto, é uma característica das células cancerosas. Esse evento celular envolve a não disjunção durante a separação dos núcleos na mitose, culminando em números alterados de cromossomos nas células resultantes.

Além dessas mutações, alterações genéticas não tão dramáticas quanto as descritas, como mutações de ponto, inserções e deleções (Capítulo 3), também caracterizam as células cancerosas.

Assim, podemos concluir que o material genético dessas células é instável, e é o alvo principal de eventos moleculares que resultam nas alterações fenotípicas.

Ainda, mecanismos que não dependem da modificação direta da sequência de nucleotídios do material genético, conhecidos como mecanismos epigenéticos, quando perturbados, podem alterar sobremaneira a atividade de genes críticos e também entram no conjunto de modificações do controle de crescimento celular normal. Exemplos de mecanismos epigenéticos são a metilação do DNA, ou seja, a introdução de grupamentos metila no DNA, principalmente em citosinas, e a alteração da organização das histonas na cromatina. Ambos os processos citados podem modificar a expressão de um gene quando acontecem em regiões importantes para sua atividade (ver Capítulo 2).

Assim, as características observadas no fenótipo das células cancerosas são provocadas por alterações genéticas e epigenéticas.

Podemos destacar como importantes propriedades de células cancerosas: a capacidade de se duplicar, independentemente dos sinais moleculares de crescimento, como a presença de fatores de crescimento; não responderem a sinais celulares de inibição do crescimento, como a inibição por contato; resistirem à morte celular programada ou apoptose; apresentarem potencial para crescimento ilimitado, não alcançando o envelhecimento celular ou senescência; induzirem a formação de novos vasos sanguíneos ou angiogênese; apresentarem invasão de tecidos vizinhos e metástase (com base em Hanahan e Weinberg, 2000). Essas características, em conjunto ou não, estão presentes em todos os tumores malignos e mostram o nível de complexidade que é marcante nessa doença.

■ O câncer e o meio ambiente: acumulando mutações

Conforme o exposto, genes críticos de controle do crescimento celular, quando alterados, podem levar ao desenvolvimento do câncer. Essas alterações podem ser herdadas, caracterizando as síndromes de câncer hereditárias, ou podem ser causadas por agentes mutagênicos aos quais todos nós temos contato diariamente.

O aparecimento de mutações nos tecidos ao longo da vida caracteriza os cânceres esporádicos, ou seja, não hereditários, e constituem a maioria dos tumores malignos.

O que são agentes carcinogênicos?

O meio que nos cerca, assim como mecanismos presentes em nossas células, produzem substâncias com o potencial de interagir e modificar o material genético. *Agentes mutagênicos* são, portanto, aqueles elementos cuja atividade promove alterações no material genético pela interação com este.

O mecanismo de ação, assim como os diferentes tipos desses elementos, estão descritos no Capítulo 3, Mutações no DNA e Mecanismo de Reparo. Na Figura 9.1 está apresentado um esquema mostrando os caminhos que um agente mutagênico pode seguir no organismo.

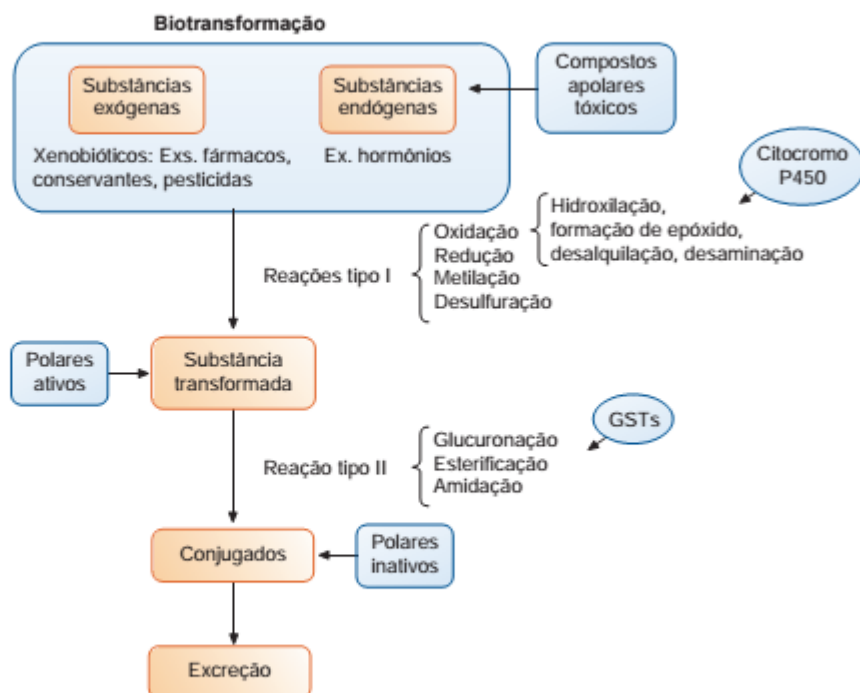



Figura 9.1 Representação do metabolismo de agentes mutagênicos no organismo.

Adutos de DNA
Segmento da molécula do DNA ligado covalentemente a um composto químico

Iniciação
Primeira etapa para a formação do câncer, na qual carcinógenos produzem modificações no material genético, tornando a célula uma célula alterada

Promoção
Segunda etapa para a formação do câncer, na qual agentes promotores como hormônios e fatores de crescimento vão promover e sustentar o crescimento das células iniciadas

Progressão
Terceira etapa para a formação do câncer, cuja célula alterada já é cancerosa e tem crescimento descontrolado

Um importante exemplo de agente mutagênico na carcinogênese química é o benzo(a)-pireno presente na fumaça de cigarro. Essa substância, após ser transformada no organismo, tem a característica de se ligar ao DNA formando adutos de DNA que podem promover a formação de mutação após a replicação do DNA. Caso os mecanismos de reparo atuantes não tenham sucesso, essas alterações são perpetuadas e, caso estejam localizadas em genes críticos para o controle do crescimento celular normal, podem se tornar carcinógenos. Portanto, nem todo agente mutagênico é um agente carcinogênico.

Para que uma substância seja considerada um carcinógeno, além de provocar alterações no equilíbrio do crescimento celular normal, este deve estar associado à produção de células cancerosas.

Os agentes mutagênicos podem ser, então, considerados iniciadores no processo de formação do câncer.

Ainda, para que a célula iniciada, ou seja, alterada por eventos mutacionais, venha a produzir um tumor maligno, esta deve ser estimulada por mecanismos que irão promover o crescimento celular e possibilitar a progressão e o desenvolvimento do câncer. Podemos, então, separar o processo da carcinogênese em 3 etapas principais: a iniciação, a promoção e a progressão.

Os agentes carcinogênicos podem ser agrupados em químicos, físicos e biológicos.

Carcinógenos químicos conhecidos são: o benzo(a)pireno, anteriormente citado, presente na fumaça de cigarro, associado ao câncer de pulmão em fumantes; a aflatoxina, presente em grãos contaminados com o fungo da espécie *Aspergillus* e associada ao câncer de fígado. Como carcinógeno físico, os raios UV, presentes na luz solar, são dos mais conhecidos e estão relacionados com o câncer de pele. O vírus do papiloma humano (HPV), associado ao câncer de colo de útero, é um exemplo de agente carcinogênico biológico.

O microambiente celular e as condições propícias para o desenvolvimento do câncer

Como vimos no item anterior, para o câncer se desenvolver, o evento inicial mais importante seria o surgimento e a perpetuação de mutações gênicas como iniciadoras da carcinogênese; entretanto, na verdade, estas não seriam suficientes para a formação e a sustentação do câncer

em desenvolvimento. Assim, para que o câncer se estabeleça em dado tecido de um organismo, deve acontecer uma adequação entre as células alteradas e o ambiente onde esta se encontra inserida. Desse modo, os fenótipos adquiridos por essas células serão perpetuados e o câncer se desenvolverá.

Microambiente celular pode ser definido como um conjunto de moléculas e células que circundam e interagem com determinada célula.

Matriz extracelular
Conjunto de elementos intercelulares que promovem a interação e a sustentação das células de um tecido

Citocinas
Moléculas existentes no meio extracelular que atuam na transmissão de sinais entre as células

No caso das células cancerosas, esse meio circundante chama-se microambiente tumoral e é composto por células normais, como células epiteliais e fibroblastos, células endoteliais, formadoras de vasos sanguíneos, e células da resposta imune. O suporte é formado pelas moléculas da matriz extracelular, como colágeno e fibronectina, e, ainda, por moléculas que fazem a comunicação entre células e que mantêm ou inibem o crescimento celular, como as citocinas. Exemplo de citocina presente no microambiente tumoral é o fator de crescimento transformador beta (TGF-β), que é um regulador do crescimento tumoral.

Existem ainda produtos do metabolismo celular e elementos como o oxigênio, presentes nesses locais, que podem alterar o meio onde a célula se encontra. Todos esses elementos devem ser propícios à manutenção do crescimento da célula cancerosa e, assim, as células cancerosas podem influenciar o ambiente que as circunda, assim como o meio circundante influencia a maneira como essas células crescem e se espalham por meio da metástase.

Célula-tronco versus célula cancerosa

Células-tronco
Células com grande potencial de replicação, não diferenciadas, que irão produzir outras células semelhantes ou células comprometidas na formação de tecidos

A questão que surge ao refletirmos sobre a condição de uma célula cancerosa é como algumas células de nosso organismo com grande capacidade de se multiplicar o fazem de maneira altamente organizada e controlada. É o caso das células-tronco que se multiplicam para formar tecidos altamente especializados.

As células-tronco são pluripotentes, ou seja, têm a potencialidade de se tornar outras células, evoluindo da maneira não especializada a uma célula diferenciada, altamente especializada na sua função.

Essas células mantêm sua capacidade replicativa durante toda a existência e estão prontas a substituírem células do tecido que tenham sido eliminadas, reconstituindo e reparando tecidos. Estas podem ser do tipo embrionárias ou adultas. As embrionárias, como o nome mesmo indica, são originadas de tecidos embrionários e podem se transformar em qualquer tipo de célula do organismo. Já as células-tronco adultas estão presentes nos tecidos, ao lado das células não tronco, e promovem a reconstituição do tecido ao qual pertencem.

Apesar de terem grande capacidade replicativa, as células-tronco crescem de maneira extremamente regulada e respondem a estímulos específicos.

Por exemplo, na linhagem de células do sangue, as células-tronco se encontram na medula óssea e se chamam células-tronco hematopoéticas e podem se transformar, ao se replicarem por estímulos de citocinas específicas, em células diferenciadas, como hemácias ou linfócitos. Já as células cancerosas, como abordado anteriormente, têm capacidade replicativa infinita e, em geral, não dependem de estímulos externos para crescer. Mas as células-tronco podem se transformar em cancerosas? Atualmente acredita-se que essas células precursoras podem adquirir alterações genéticas e epigenéticas, sendo que, nas síndromes de câncer hereditárias (ver adiante), como nas famílias que têm mutação herdada, essas células já apresentariam predisposição aumentada a se tornarem cancerosas. Acredita-se que o tipo de câncer originado por essas células precursoras têm características mais agressivas.

■ Oncogenes e supressores tumorais

Até aqui, descrevemos as principais características do desenvolvimento do câncer. Agora, vamos descrever os principais atores nesse cenário. Duas grandes classes de genes, cujos

produtos proteicos estão envolvidos no controle do crescimento celular normal, são descritos pelos pesquisadores como críticos no processo de desenvolvimento do câncer: os oncogenes e os supressores tumorais. Os genes ligados aos processos de reparo do DNA também são importantes no processo, no entanto, não trataremos desses tipos de genes neste capítulo.

Oncogene
Classe de genes caracterizados pela presença de mutação ativadora ou expressão aumentada, associada ao processo de formação do câncer

Sarcoma
Tipo de câncer originado de células do tecido conjuntivo

Proto-oncogene
Classe de genes associados ao controle positivo do ciclo celular que, quando apresentam mutação ou têm sua expressão aumentada, se transformam nos oncogenes

O que são oncogenes?

A descrição da classe de genes denominados oncogenes se deve à descoberta de um tipo de vírus contendo RNA como material genético, os retrovírus, que poderiam causar transformação celular e a formação de tumores malignos. A primeira descrição foi feita em 1910 pelo pesquisador Peyton Rous, que mostrou que a presença de vírus induzia à formação de sarcoma em galinhas. Com as modernas técnicas de genética molecular e bioquímica foi possível demonstrar os principais mecanismos moleculares e a descrição dos principais genes envolvidos no processo. Posteriormente, ficou claro para os pesquisadores que muitos dos oncogenes se tratavam de genes presentes normalmente nas células e que tinham função definida, associada sempre ao controle positivo do crescimento celular. Assim, esses genes presentes no genoma da célula normal foram denominados proto-oncogenes. Estes, por sua vez, apresentam uma contraparte viral, ou seja, genes celulares homólogos aos virais. Para diferenciar a origem do oncogene, dá-se o nome *v-onc*, para origem viral, e *c-onc*, para origem celular.

Muitos exemplos importantes de oncogenes e seus produtos foram descritos.

A característica que mais chama atenção nos oncogenes é a propriedade de ativar o crescimento celular, fazendo, portanto, o controle positivo exacerbado.

Para tanto, os proto-oncogenes adquirem mutação ativadora, o que caracteriza ganho de função para seu produto proteico. Essa ativação pode ser resultado de uma mutação de ponto ou por amplificação gênica ou, ainda, por translocação cromossômica e formação de um gene quimérico. Exemplos importantes são: gene c-*ras*, mutação de ponto ativadora e câncer de pulmão; amplificação do gene *c-myc* e neuroblastoma; amplificação de *c-erbb2* e câncer de mama; e, por fim *bcr/c-abl* e formação do cromossomo Filadélfia por translocação cromossômica na leucemia mieloide crônica que iremos descrever a seguir.

Enzimas
Proteínas ou RNAs com características catalíticas, ou seja, que promovem reações químicas em sistemas vivos

Quinase
Enzima de natureza proteica que promovem a introdução de grupamentos fosfato em componentes celulares

Mitógeno
Substância que estimula a proliferação celular

A proteína Ras

As proteínas Ras constituem uma família de proteínas, de pequeno tamanho molecular, do tipo trifosfatase de guanosina ou GTPase, que faz parte de via de sinalização ativadora de uma via de eventos moleculares que vão impulsionar a divisão celular. Vias desse tipo são designadas cascatas e são formadas essencialmente por proteínas do tipo enzimas que, ao serem alteradas, irão alterar a próxima proteína, em um efeito do tipo cascata. Isso ocorre principalmente pela introdução e a retirada de grupamentos fosfato ou seja, fosforilação e defosforilação. Essa via é conhecida como via da MAP quinase (proteinoquinase ativada por mitógeno) e é uma importante via de controle da proliferação celular, cujos componentes envolvem vários produtos de oncogenes.

A Ras tem o importante papel de regular a transmissão de sinais provenientes do meio externo, por meio de fatores de crescimento e sua ligação aos receptores específicos, presentes na membrana celular.

Na Figura 9.2, você poderá encontrar a representação dessa importante via de controle da proliferação celular.

Receptores celulares
Moléculas de natureza essencialmente proteica que interagem com substâncias externas às células, desencadeando alterações nestas

Alguns exemplos de fatores de crescimento são o fator de crescimento epidérmico (EGF), que controla a proliferação e a diferenciação de células epiteliais, e o fator de crescimento derivado de plaquetas (PDGF), que estimula a proliferação de células do tecido conjuntivo. Os receptores celulares, por exemplo o receptor para fator de crescimento epidérmico (EGFR), são proteínas transmembranares, ou seja, com uma porção exterior à célula, na qual o fator de crescimento se liga, e uma porção interna, com atividade de proteinoquinase. No caso dos exemplos citados, estes são do tipo tirosina quinase.

Figura 9.2 Papel das proteínas Ras e Ras mutante na via da MAP quinase.

A seguir, iremos descrever a cascata de eventos da via MAP quinase pela ativação do receptor do fator de crescimento epidérmico (EGFR). A tirosina quinase acoplada ao EGF é ativada após a sua ligação ao EGFR, que está acoplado a uma tirosina quinase. A ligação do EGF ao EGFR resulta na ligação de duas moléculas da quinase, formando dímeros, e a ativação desta, que se dá pela fosforilação de tirosinas fora do domínio catalítico. Assim, cada componente do dímero fosforila o seu par. Após a ativação da quinase, a transmissão do sinal ocorre pela fosforilação de proteínas citoplasmáticas e o recrutamento de proteínas adaptadoras que se ligam especificamente a resíduos de tirosinas fosforiladas tanto no EGFR como em outros substratos do EGFR. Na última etapa da cascata de sinalização, o sinal é transmitido para o núcleo da célula, ativando fatores de transcrição específicos que levam à proliferação celular. A proteína Ras está associada ao receptor e está ativa quando ligada a uma molécula de GTP (guanosina trifosfato) e inativa quando ligada à GDP (guanosina difosfato). Assim, quando ligada à GTP, ela participa como transmissora de sinais para a fosforilação das proteínas da cascata.

Nos cânceres humanos, o gene *ras* aparece frequentemente mutado, sendo um dos oncogenes mais presentes no processo de formação do câncer.

Mutação de ponto
Tipo de mutação que ocorre na molécula de DNA em que um nucleotídio é trocado por outro nucleotídio diferente

A principal mutação é uma mutação de ponto, ativadora, existente no códon 12, havendo na proteína uma substituição do aminoácido valina para glicina. Essa alteração leva a uma ativação constitutiva da Ras, que permanece na estrutura associada ao GTP e, portanto, possibilita a transmissão de sinais e a ativação da cascata de maneira independente da presença de fatores de crescimento. Assim, a via da MAP quinase, ativa, aumenta a proliferação celular, resultando em carcinogênese. Outras mutações pontuais, com troca de aminoácidos, são descritas nas posições dos códons 59 e 61 do gene *ras*, todas levando à ativação constitutiva. O câncer de pulmão é um tipo no qual essa alteração genética é bastante frequente. Outros tipos de tumores como o câncer de cólon, próstata e bexiga também apresentam esse tipo de mutação.

A amplificação gênica: c-myc e c-erbb2

Amplificação gênica
Tipo de mutação que leva à ocorrência de múltiplas cópias de um gene em um genoma

Causadas por eventos mutacionais não compreendidos, alguns tumores malignos podem ter várias cópias, por amplificação gênica, de genes críticos para a proliferação celular. A proteína myc é um fator de transcrição que age aumentando a produção das mais diversas proteínas, inclusive com funções ativadoras do ciclo celular. Seu gene, o *c-myc*, é um conhecido oncogene ligado a vias que levam à proliferação celular. A presença de cópias desse gene, além dos 2

Alelo
Maneira alternativa de um gene que ocupa um *locus* cromossômico específico

alelos normais, acaba promovendo um maior crescimento celular, que, junto a outras alterações gênicas, podem levar à formação do câncer.

Um tumor raro, do tipo embrionário, chamado de neuroblastoma, está associado a esse tipo de alteração.

Já a proteína erbb2 tem função de receptor celular para fator de crescimento epidérmico humano (EGF), descrito anteriormente.

O *c-erbb2*, se encontra amplificado em aproximadamente 27% dos cânceres avançados de mama.

Como receptor, a erbb2 tem uma porção de proteína exterior à célula e uma porção com função de fosforilação que ativa uma série de proteínas que formam vias de ativação do crescimento celular. A detecção dessa alteração implica o diagnóstico e a terapêutica. Anticorpos e inibidores químicos específicos vêm sendo desenvolvidos para bloquear o excesso de erbb2.

O cromossomo Filadélfia e a leucemia mieloide crônica

Translocação cromossômica
Rearranjo de segmentos de cromossomos não homólogos

Cromossomo Filadélfia
Cromossomo 22 anormalmente curto em virtude da translocação cromossômica que envolve os cromossomos 9 e 22 presentes na leucemia mieloide crônica

Na leucemia mieloide crônica (CML) encontramos um exemplo de translocação cromossômica que causa a formação de um oncogene decisivo no desenvolvimento desse tipo de câncer. Essa translocação dá origem ao que é conhecido como cromossomo Filadélfia, e ocorre entre o cromossomo 9, na região do gene da Abl (uma tirosina quinase), e o cromossomo 22, na região do gene da Bcr. A união desses 2 genes (*Bcr/c-Abl*) resulta na produção de uma proteína de fusão oncogênica: a Bcr/Abl.

A Bcr/Abl tem uma atividade catalítica maior do que a c-Abl porque nessa translocação perde-se o domínio da c-Abl responsável pela autorregulação da atividade catalítica da c-Abl.

Assim, a desregulação da c-Abl leva à fosforilação de diversas proteínas, o que resulta na regulação positiva do ciclo celular e, portanto, no crescimento das células afetadas pela mutação.

O cromossomo Filadélfia foi descrito por pesquisadores em 1960 e o nome se deve à cidade dos Estados Unidos onde foi feita a descoberta. A detecção desse tipo de aberração cromossômica é de suma importância para o diagnóstico, já que em torno de 90% dos pacientes de CML a mutação está presente. A Figura 9.3 mostra o ponto de translocação e a formação do *cromossomo Filadélfia*.

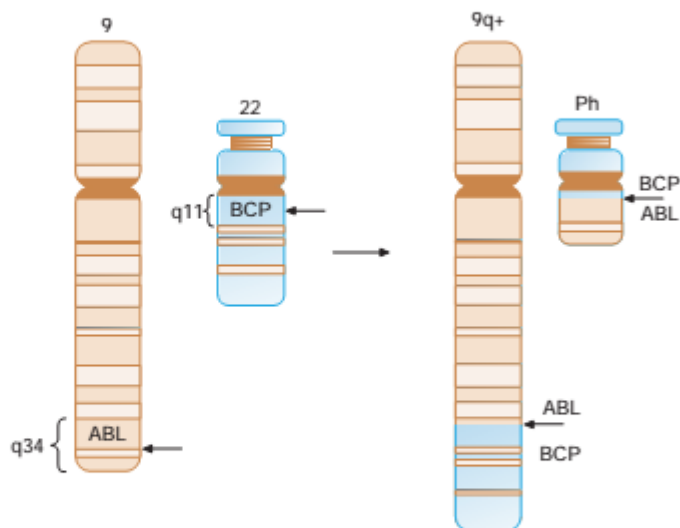



Figura 9.3 Representação do cromossomo Filadélfia e do ponto de quebra.

O que são genes supressores de tumor?

Como foi descrito, os oncogenes surgem por mutações que levam à exarcebação da função de seu produto proteico. Já os genes conhecidos como supressores tumorais são caracterizados pela perda da função de seu produto proteico, principalmente por mutações deletérias.

Genes supressores tumorais
São genes cujos produtos proteicos participam do controle negativo do ciclo celular

As proteínas supressoras tumorais estão associadas ao controle negativo do crescimento celular, promovendo, em geral, uma diminuição da progressão do ciclo celular.

Não é de se espantar que esses genes estão entre os mais frequentemente alterados nos cânceres. A seguir, apresentaremos 2 genes supressores tumorais que podem ser considerados os mais importantes representantes dessa classe de genes: os genes *Rb* e *TP53*.

A pRb e o retinoblastoma

Retinoblastoma
Câncer que acomete a retina

Deleção gênica
Tipo de mutação no qual um segmento de DNA contendo um ou mais genes são perdidos

O retinoblastoma é um tipo de câncer que acomete a retina e é causado por alterações ou pela ausência do gene *Rb* por deleção gênica. O *Rb* foi o primeiro gene supressor tumoral a ser clonado.

O estudo das alterações nesse gene tem importância na história da pesquisa em câncer, pois foi por meio desse exemplo de alteração molecular que Knudson, um pesquisador envolvido com o desenvolvimento desse câncer em crianças, em publicação de 1971, propôs a hipótese dos 2 eventos que levou ao entendimento do processo de formação do retinoblastoma e trouxe luz ao conhecimento das implicações genéticas no câncer. Nessa hipótese, Knudson propõe que, no retinoblastoma familiar, o indivíduo portador de alteração em um dos alelos do gene Rb deve necessariamente desenvolver alteração deletéria no outro alelo para que o câncer na retina surgisse. Assim, apenas um evento mutacional não seria suficiente para a formação do câncer. A Figura 9.4 mostra um esquema que ilustra a Hipótese de Knudson.

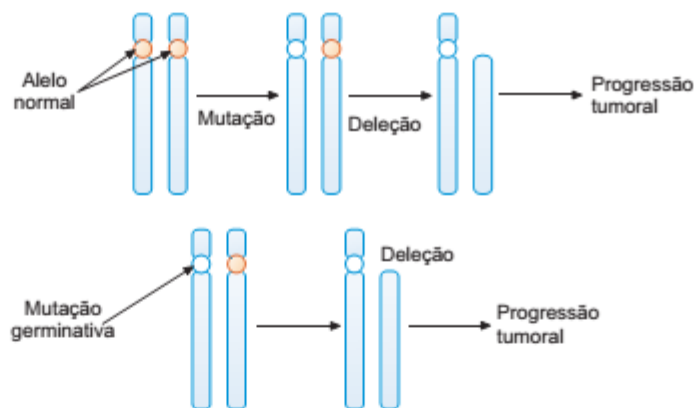



Figura 9.4 Representação da Hipótese dos 2 eventos de Knudson.

Fator de transcrição
Proteínas que se ligam a sítios específicos no genoma e que ativam a transcrição

DNA polimerase
Enzima com atividade de produzir novas moléculas de DNA com base em uma molécula de DNA preexistente utilizada como molde

A Rb é uma importante proteína que age controlando negativamente o ciclo celular. Essa proteína quando não está sob a forma fosforilada, encontra-se associada a um fator de transcrição chamado de E2F, que ativa a produção de diversas proteínas que, por sua vez, irão promover a progressão do ciclo celular. Um exemplo dessas proteínas reguladas por E2F é a própria DNA polimerase, a enzima que sintetiza novas moléculas de DNA (Capítulo 1). Assim, bloqueando o E2F, a Rb bloqueia o ciclo celular que só seguirá para a etapa de síntese (fase S) quando esta receber grupamentos fosfato pelas enzimas chamadas de quinases, que são específicas e atuantes no ciclo celular. Ao ser fosforilada, a Rb se desliga da E2F.

Assim, a Rb controla negativamente o ciclo celular e tem sua atividade biológica prejudicada por alterações moleculares em células cancerosas, sendo considerada, portanto, uma proteína supressora tumoral.

P53, a guardiã do genoma

Dentre as várias proteínas supressoras tumorais descobertas nas últimas décadas, a p53 tem se destacado por ser uma proteína central na integração de diversas vias que regulam a vida e a morte celular após situações de estresse celular, como a presença de lesões no DNA. O papel

central na manutenção da integridade do genoma levou alguns pesquisadores a chamarem a p53 de "a guardiã do genoma".

A p53 atua como *fator de transcrição* e faz seu papel de mediadora da supressão tumoral, ativando a expressão de proteínas que podem levar à parada do ciclo celular, à apoptose, ou, ainda, a ativar vias de reparo de DNA.

A ativação da p53 se dá por modificações de sua estrutura pela adição de grupamentos fosfato, por meio da fosforilação por enzimas quinases que, por sua vez, são ativadas quando a célula está sob estresse. As alterações estruturais produzidas por essas modificações na molécula levam à sua estabilização. Normalmente, a p53 é bastante lábil com uma meia-vida curta de aproximadamente 15 min. Um mecanismo eficiente de degradação, pela proteína MDM2, leva à destruição da p53, e é impedido de acontecer quando a p53 está sob a forma ativa. A Figura 9.5 mostra as principais vias de ativação da p53 e seus principais alvos.

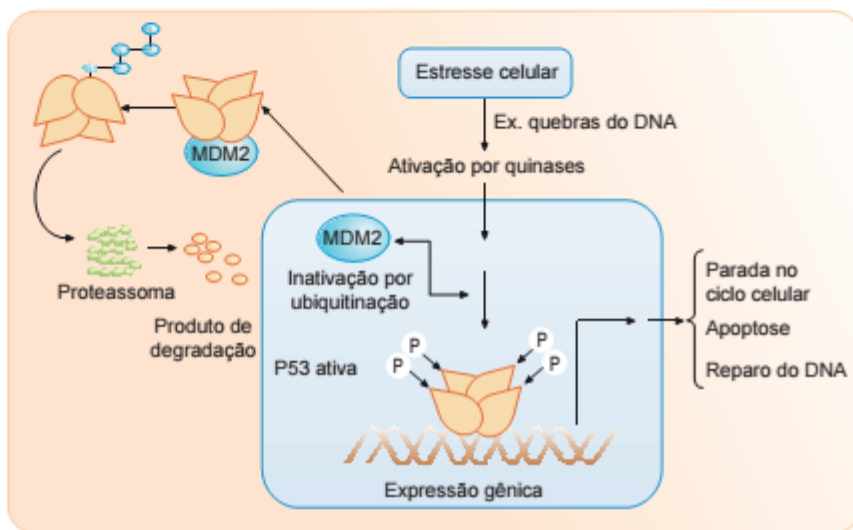



Figura 9.5 Esquema representando as principais vias de ativação da p53 e seus principais alvos.

Mutações no gene da p53, o *TP53*, estão presentes em mais de 50% de todos os cânceres e alteram a atividade normal da proteína. Na verdade, as mutações no gene *TP53* são os eventos mutacionais mais comuns no câncer humano.

O tipo de mutação mais observado no *TP53* são mutações de ponto, com troca de nucleotídios e de códons, que modificam a estrutura proteica.

Essas alterações são encontradas majoritariamente no domínio central da proteína, e é onde a ligação ao DNA acontece, sendo fundamental para sua atividade como fator de transcrição. Curiosamente, as mutações podem anular a atividade da p53 mesmo estando localizadas em apenas um alelo, por meio de um mecanismo conhecido como dominante negativo. Ainda, algumas dessas mutações podem alterar a atividade da p53 de maneira a promover novas atividades biológicas, tornando-a uma proteína oncogênica. Portanto, além de impedir o seu papel de supressora tumoral, algumas mutações ainda facilitam o processo de evolução do câncer. O estudo com camundongos *knockout* tem sido muito importante na compreensão dos mecanismos da carcinogênese associados a supressores tumorais. O desenvolvimento desses animais é feito por meio de técnicas moleculares que desligam determinado gene do genoma em células-tronco embrionárias do animal, tendo como consequência, posteriormente, um animal adulto que não expressa a proteína-alvo. De modo interessante, animais *knockout* para a p53 se desenvolvem e crescem normalmente, no entanto, esses camundongos têm grande predisposição a desenvolver tumores durante a sua vida.

Camundongos *knockout*
Linhagem de camundongo gerada pela anulação de determinado gene ou grupo de genes por técnicas de biologia celular e molecular

Ainda, como a proteína do retinoblastoma descrita anteriormente, mutações no gene *TP53* podem ser herdadas e levar ao desenvolvimento de uma síndrome rara, conhecida como síndrome de Li-Fraumeni (ver adiante), cujas famílias são caracterizadas pela presença de um grande número de indivíduos afetados por diferentes tipos de tumores. Assim, o impedimento da atividade da p53 é um dos eventos mais importantes na carcinogênese humana.

■ Síndromes do câncer hereditário ou familial

A grande maioria dos tumores malignos é considerada esporádica, ou seja, são cânceres que se desenvolvem durante a vida do indivíduo, em múltiplas etapas, podendo levar muitos anos até o seu estabelecimento. Já o câncer familial está associado a componentes genéticos herdados, levando à maior predisposição ao desenvolvimento de tumores, os quais surgem notadamente em indivíduos mais jovens. Outra característica importante que aponta para esse tipo de doença é a presença de bilateralidade, como observado no retinoblastoma herdado, quando os 2 globos oculares do indivíduo são afetados. Também no câncer de mama hereditário, as duas mamas podem ser afetadas. Portanto, mutações esporádicas, ou seja, mutações adquiridas, caracterizam os cânceres esporádicos, e mutações hereditárias caracterizam os cânceres hereditários e estão presentes em todas as células do indivíduo portador. As mutações germinativas são aquelas presentes nas células de linhagem germinativa (óvulos e espermatozoides) que serão, portanto, transmitidas aos descendentes. As mutações herdadas têm origem nas mutações germinativas.

Mutação esporádica
Mutação que surge nas células somáticas, sendo adquiridas durante o período de vida do indivíduo

Mutação hereditária
Mutação adquirida pela herança genética

Mutação germinativa
Mutação presente na linhagem germinativa e que é transmitida aos descendentes do indivíduo afetado

Apesar de síndromes de câncer hereditário não serem frequentes, o enorme peso que os familiares carregam por vivenciarem o aparecimento de vários casos as tornam de fundamental importância para a saúde pública. Graças ao aconselhamento genético fornecido por profissionais especializados, bem como ao desenvolvimento de novos métodos de diagnóstico por imagem ou por meio de técnicas de biologia molecular, essas famílias podem ser assistidas, por meio do diagnóstico precoce, e evitar o pior. A Tabela 9.1 apresenta alguns exemplos de síndromes de câncer familial.

■ **Tabela 9.1** Síndromes de câncer familial.

Síndrome	Gene	*Locus*	Tipos principais de câncer	OMIM
Câncer de mama familial 1	*BRCA1*	17q21	Câncer de mama e ovário	604370
Câncer de mama familial 2	*BRCA2*	13q12	Câncer de mama, também em homens, e ovário	612555
Síndrome de Li-Fraumeni 1	*TP53*	17p13	Sarcomas, tumores cerebrais, câncer de mama, leucemia	151623
Polipose adenomatosa familial	*APC*	5q21	Câncer colorretal, câncer duodenal e gástrico	175100
Adenomatose endócrina múltipla 2 (MEN2 A e B)	*RET*	10q11	Câncer de tireoide do tipo medular, adenomas da paratireoide, feocromocitoma	171400, 162300
Retinoblastoma	*RB1*	13q14	Retinoblastoma	180200
Síndrome de Lynch I	*MSH2*	2p22	Câncer colorretal	120435
Neurofibromatose tipo1	*NF1*	17q11	Neurofibromas	162200
Neurofibromatose tipo2	*NF2*	22q12	Schwannomas, meningiomas e ependimomas	101000

A síndrome do câncer de mama familial

De todos os casos de câncer de mama, 5% a 10% são de origem familial. Destes, menos de 25% têm origem em mutações presentes em genes conhecidos. Os mais importantes genes envolvidos com o câncer de mama hereditário são: *BRCA1*, *BRCA2* e *TP53*. Outros genes também têm sido descritos, como: *ATM*, *PTEN*, *CHEK2* e *PALB2*. Assim, a grande maioria dos casos de câncer de mama familial não tem origem genética definida e pode ser explicada pelos polimorfismos presentes em genes cujos produtos participam de vias de metabolismo de carcinógenos e hormônios, em vias de reparo e/ou de controle do ciclo celular. Essas variantes gênicas modificam o risco de apresentar câncer de mama e atuam em conjunto, constituindo um tipo de herança chamado de herança poligênica para o câncer de mama familial.

Herança poligênica
Tipo de herança em que o fenótipo é resultado da combinação da expressão de variantes de um conjunto de genes

A chamada síndrome de câncer de mama e ovário, que pode ser do tipo 1 e do tipo 2, é uma síndrome hereditária autossômica dominante e está associada à presença de mutações herdadas nos genes *BRCA1* e *BRCA2*, respectivamente. Acredita-se que, aproximadamente, 50% das famílias com esse tipo de doença têm mutação em *BRCA1*. Esse gene, localizado no cromossomo 17, é classificado como supressor tumoral e seu produto, a proteína BRCA1, atua

em vias de reparo. No entanto, nem todas as suas funções estão definidas. Os pesquisadores identificaram mais de 1.000 mutações, sendo que apenas parte delas seria responsável pela alteração da função supressora tumoral da proteína. Ainda, diferentes mutações conferem riscos diferentes, por exemplo, as famílias de origem judia, pertencentes ao grupo Askenazi, podem apresentar duas variantes genéticas: as mutações 187delAG e 5385insC, que promovem um risco de câncer de mama e ovário nas famílias de 56% e 16%, respectivamente, aos 70 anos. Outros tipos de câncer também podem estar associados nessas famílias, como o câncer de próstata. Como o *BRCA1*, o gene *BRCA2*, que está localizado no cromossomo 13, é classificado como supressor tumoral. Seu produto, a proteína BRCA2, atua em vias de reparo e de controle do crescimento celular. Pesquisadores identificaram seu papel na fase de separação das células durante o crescimento celular, no entanto nem todas as suas funções estão definidas. A presença de mutações também está associada ao aparecimento precoce do câncer de mama e ovário, mas está principalmente associada ao câncer de mama em homens. Os pesquisadores identificaram mais de 800 mutações nesse gene. Outra doença ligada à alteração em BRCA2 é a Anemia de Fanconi tipo D1, que ocorre quando a proteína é quase inexistente nas células por defeito no gene. Essa doença está relacionada ao aparecimento de inúmeros cânceres na família portadora da alteração, principalmente leucemias.

A síndrome de Li-Fraumeni

Essa síndrome hereditária é autossômica dominante, rara, e está associada à presença de vários tipos de câncer em idade precoce. Os principais tipos de câncer encontrados nessa síndrome são: câncer de mama, sarcomas, câncer de cérebro, câncer corticoadrenal. As mutações no gene *TP53* são causadoras da síndrome Li-Fraumeni tipo 1, enquanto as do tipo 2 são causadas por mutações no gene *CHEK2*. O gene *TP53*, descrito anteriormente, e a proteína p53 têm funções fundamentais no controle do crescimento celular normal e na manutenção da estabilidade do genoma. A presença de mutações germinativas da p53 está associada à maioria das famílias Li-Fraumeni. Algumas não apresentam todas as características de Li-Fraumeni e são classificadas como Li-Fraumeni-like; nesses 2 tipos, o câncer de mama é o mais frequente. Por volta de 400 mutações germinativas foram identificadas pelos pesquisadores.

Câncer de cólon não polipose hereditário e polipose adenomatose familial

O câncer de cólon não polipose hereditário (sigla em inglês HNPCC) também é conhecido como síndrome de Lynch e pode ser do tipo 1 ou do tipo 2. Aproximadamente 5% dos casos de câncer colorretal são originados por essa síndrome que se caracteriza pela presença específica de tumores do cólon, tipo 1, ou pela presença de tumores extracólon, como no estômago, no endométrio, na vesícula biliar e no pâncreas. Os tumores que surgem nos indivíduos portadores apresentam instabilidade no genoma característica. Os genes envolvidos nessa síndrome são genes relacionados com os mecanismos de reparo de pares de bases mal pareadas, como o gene MSH2, cuja presença de mutações em heterozigose leva ao desenvolvimento da síndrome de Lynch do tipo 1. Já a polipose adenomatose familial está ligada à presença de mutações no gene *APC* e se caracteriza pela presença de muitos pequenos pólipos na região do cólon e do reto. A APC é uma proteína multifuncional com papel de supressora tumoral. Mutações deletérias levam ao desenvolvimento do câncer colorretal, duodenal e gástrico.

Síndromes da neoplasia endócrina múltipla (NEM)

As síndromes da neoplasia endócrina múltipla do tipo 1 e do tipo 2 (NEM 1 e 2) são síndromes genéticas raras que acometem principalmente o sistema de glândulas paratireoides, tireoides e adrenais, com o surgimento de tumores. Na NEM 1, a maioria dos tumores é benigna, sendo que alguns podem evoluir para a malignidade, sendo a mais frequente causa de morte dos portadores. Mutações deletérias no gene *MEN 1*, um supressor tumoral, está associada a

essa doença. A NEM 2 pode ser subdividida em NEM 2A e NEM 2B, de acordo com o fenótipo apresentado pelos portadores da doença. A NEM 2 está associada a mutações ativadoras no proto-oncogene *RET*. Esse gene codifica um receptor celular do tipo tirosina quinase, que fica permanentemente ativado quando alterado por mutações de ponto, herdadas. Assim, a proteína mutante permanece enviando sinais de ativação para a cadeia da via da MAP quinase, o que acaba por desencadear a formação de tumores, à maneira da proteína Ras ativada, discutida anteriormente. De maneira interessante, de acordo com o tipo de mutação ativadora em *RET*, o fenótipo da doença é diferente. Assim, a mutação no códon 634 está associada à NEM 2A, que se caracteriza pela presença, além de câncer medular da tireoide, de feocromocitoma, que é um tipo de tumor geralmente benigno da adrenal, e hiperparatireoidismo. Já a mutação no códon 918 está associada à NEM 2B, que também apresenta, caracteristicamente, câncer medular da tireoide, mas também feocromocitoma e neuromas, que são tumores benignos do tecido nervoso que se caracterizam por uma degeneração lenta deste.

Neurofibromatose do tipo 1 e do tipo 2

O conjunto de doenças chamadas de neurofibromatoses está associado ao aparecimento de tumores benignos e malignos, principalmente cutâneos e no sistema nervoso. A presença de mutação em 2 genes *NF1* e *NF2* determina se a neurofibromatose é do tipo 1, no caso de mutação presente em NF1 ou do tipo 2 em NF2. A neurofibromatose do tipo 1, também conhecida como síndrome Von Recklinghausen, é caracterizada pela presença de manchas na pele de cor café com leite e de neurofibromas, que são pequenos tumores benignos do sistema nervoso periférico. O gene *NF1* codifica a proteína neurofibromina e mutações deletérias suprimem a via da Ras (ver anteriormente). Já a neurofibromatose do tipo 2 é uma doença bem mais rara do que a neurofibromatose do tipo 1. Acredita-se que 1 entre 40.000 indivíduos seja portador dessa doença. Ela se caracteriza pela presença de tumores nas células nervosas, como nas células de Schwann (schwannomas), nas células meníngeas (meningiomas) e nas células gliais (ependimomas). Tumores acústicos, ou seja, localizados nos ouvidos, estão associados a essa doença. O gene *NF2* codifica a proteína chamada de merlina, que é uma proteína do citoesqueleto, associada à adesão célula a célula em vários tecidos, inclusive o tecido nervoso.

G·F Genética em Foco

Teste genético e aconselhamento em câncer

O objetivo do aconselhamento genético em câncer é o mesmo para as outras doenças genéticas, ou seja, procurar esclarecer as consequências e problemas com a ocorrência ou o risco de ocorrência de um câncer de origem genética em uma família. Esse processo é realizado por profissionais treinados, quais sejam, biólogos e médicos geneticistas, enfermeiros e psicólogos. A procura de aconselhamento genético em câncer deve ser feita quando o indivíduo apresenta uma ou mais das seguintes características: câncer quando jovem, 2 ou mais tipos de câncer ou diversos casos de câncer na família. Com base na avaliação da história pregressa de câncer na família e no próprio indivíduo, os profissionais irão construir heredogramas e calcular o possível risco de aparecimento da doença. Ao mesmo tempo, o indivíduo será instruído a fazer exames clínicos de acompanhamento e testes genéticos para ele e seus familiares. Os testes genéticos são realizados com o intuito de identificar a alteração gênica nos indivíduos e familiares e são indicados apenas após o diagnóstico definido como pertencente a uma das síndromes de câncer. As síndromes mais indicadas para o teste genético são: câncer de mama familial do tipo 1 e 2; polipose adenomatosa familial; adenomatose endócrina múltipla 2A e B e Li-Fraumeni, entre outras. A análise molecular de genes-alvo por sequenciamento de bases é o tipo de teste mais utilizado para a identificação de mutação herdada em câncer.

RESUMO

- O câncer é um conjunto de doenças caracterizadas por crescimento celular descontrolado, morfologia do tecido alterada e capacidade de migração das células para outros órgãos
- O material genético das células cancerosas é caracteristicamente instável, e é alvo principal de eventos moleculares que resultam nas alterações fenotípicas
- Além das mutações presentes no DNA das células malignas, genes importantes no controle da proliferação celular podem ter sua expressão alterada por mecanismos epigenéticos, ou seja, processos moleculares que modificam a expressão de genes sem alterar sua sequência de nucleotídios. Dentre esses mecanismos estão a metilação de regiões controladoras da expressão de um gene e a organização das histonas na cromatina
- Pode-se observar que as características mais comuns das células cancerosas são: crescimento independente da presença de fatores de crescimento; não apresentarem inibição por contato; resistência à apoptose; crescimento ilimitados sem apresentarem senescência; indução da angiogênese; metástase
- As mutações são adquiridas ou herdadas, sendo que as primeiras caracterizam o câncer esporádico, ou seja, não hereditário, que é, de longe, o mais frequente
- Conceituam-se agentes mutagênicos como aquelas substâncias com o potencial de interagir e modificar o material genético. Um agente mutagênico pode se tornar um agente carcinogênico
- Os agentes mutagênicos podem, então, iniciar o processo de formação do tumor maligno. Após serem iniciadas, as células devem ter seu crescimento sustentado e, então, progredir para a formação do câncer. Assim os agentes mutagênicos se tornam agentes carcinogênicos
- Os agentes carcinogênicos podem ser agrupados em agentes químicos, físicos e biológicos, de acordo com sua origem. O benzo(a)pireno e o câncer de pulmão; raios UV e câncer de pele; e o vírus do papiloma humano e câncer de colo de útero são exemplos dessas classes de substâncias
- Microambiente celular é o conjunto de moléculas e células que circundam e interagem com determinada célula e, no câncer, esse microambiente tumoral promoverá e sustentará o crescimento da célula maligna
- Duas classes de genes são críticas no processo de desenvolvimento do câncer: os oncogenes e os supressores tumorais
- Os oncogenes são genes que se caracterizam pela presença de mutação ativadora ou expressão aumentada, associada ao processo de formação do câncer. Muitos oncogenes podem ter origem viral, assim, foi adotada pelos pesquisadores a denominação *v-onc*, para origem viral, ou *c-onc*, para origem celular
- Exemplos importantes de oncogenes são: gene c-*ras*, mutação de ponto ativadora e câncer de pulmão; amplificação dos genes *c-myc* e neuroblastoma e *c-erbb2* e câncer de mama; *bcr/c-abl;* e a formação do cromossomo Filadélfia por translocação cromossômica na leucemia mieloide crônica
- Uma importante via de controle da proliferação celular é a conhecida via da MAP quinase, cujos componentes envolvem vários produtos de oncogenes como a Ras e a Ret
- Os genes supressores tumorais são genes cujos produtos proteicos controlam negativamente o ciclo celular. Sua atividade se encontra impedida nas células cancerosas
- Exemplos de importantes supressores tumorais são o *Rb* que produz a proteína retinoblastoma e o *TP53* que produz a proteína p53
- As mutações em oncogenes e supressores tumorais, assim como em outras classes de genes como os que participam dos mecanismos de reparo do DNA, podem ser herdadas. Essas alterações hereditárias, ou seja, passadas de parental para a prole, vão caracterizar as síndromes do câncer hereditário
- As mutações esporádicas, ou seja, mutações adquiridas, caracterizam os cânceres esporádicos, e mutações hereditárias caracterizam os cânceres hereditários e estão presentes em todas as células do indivíduo portador
- As mutações germinativas são as mutações presentes nas células de linhagem germinativa (óvulos e espermatozoides) que serão, portanto, transmitidas aos descendentes. As mutações germinativas vão originar as mutações herdadas que, por sua vez, vão produzir as síndromes do câncer hereditário
- Alguns exemplos de síndromes do câncer hereditário são o câncer de mama familial, cujos genes *BRCA1* e *BRCA2* estão implicados; a síndrome de Li-Fraumeni, cujo gene envolvido é o *TP53*, e o retinoblastoma, com o *RB1*, entre outros.

AUTOAVALIAÇÃO

9.1 Quais são as principais características biológicas de uma célula cancerosa?

9.2 Conceitue mutação gênica e descreva sua importância no processo de formação do câncer.

9.3 O que são mecanismos epigenéticos?

9.4 Comente as principais características fenotípicas que uma célula adquire no caminho para a formação do tumor maligno.

9.5 Como você explica que uma célula com grande capacidade replicativa, como a célula-tronco, não tenha as mesmas propriedades das células cancerosas?

9.6 Diferencie mutagênese de carcinogênese.

9.7 O que são adutos de DNA?

9.8 Dê exemplos de agentes carcinogênicos químicos, físicos e biológicos.

9.9 Conceitue oncogenes e genes supressores tumorais.

9.10 Diferencie proto-oncogene de oncogene.

9.11 Descreva, em linhas gerais, o papel da proteína Ras e sua forma ativada.

9.12 O que é o cromossomo Filadélfia?

9.13 Descreva o papel do gene *Rb* como supressor tumoral.

9.14 Comente a afirmativa: "A proteína p53 é considerada a guardiã do genoma de uma célula.".

9.15 Quais são os principais efeitos observados na proteína p53 resultantes da presença de mutações no seu gene, o *TP53*?

9.16 Diferencie mutação esporádica, mutação hereditária e mutação germinativa.

9.17 Na sua opinião, qual é o impacto para uma família ao conhecer que um de seus componentes é portador de mutação germinativa em um gene supressor tumoral?

9.18 Cite 3 síndromes do câncer familial e o gene envolvido.

BIBLIOGRAFIA

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. *Biologia molecular da célula*. 5ª ed., Porto Alegre: Artmed, 2010.

AMERICA BOARD OF GENETIC COUNSELING (ABCG). Disponível em: http://www.abgc.net/.

FORD, D.; EASTON, D.F.; STRATTON, M. *et al.* Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* 1998; 62:676-89.

HANAHAN, D.; WEINBERG, R.A. 2000. The hallmarks of cancer. *Cell* 100:57-70.

HERCEG, Z. Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. *Mutagenesis* 2007; 22:91-103.

ONLINE MENDELIAN INHERITANCE IN MAN (OMIM) – http://www.ncbi.nlm.nih.gov/omim/.

PETITJEAN, A.; MATHE, E.; KATO, S. *et al.* Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum. Mutat* 2007; 28:622-29.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética*. 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

STRACHAN, T.; READ, A.P. *Genética molecular humana*. 2ª ed., Porto Alegre: Artmed, 2002.

STRATTON, M.R.; CAMPBELL, P.J.; FUTREAL, P.A. The cancer genome. *Nature* 2009; 458:719-24.

VOGELSTEIN, B.; KINZLER, K. *The genetic basis of human cancer*. 2ª ed., EUA: McGraw-Hill Professional, 2002.

WOOD, L.; PARSONS, D.; JONES, S. *et al.* 2007. The genomic landscapes of human breast and colorectal cancers. *Science*, 318:1108-113.

YOUNG, I. D. *Genética médica*. Rio de Janeiro: Guanabara Koogan, 2007.

10

Tecnologias Moleculares e suas Aplicações

Objetivos de estudo

Compreender a importância do desenvolvimento em técnicas de análise molecular para a genética atual

Compreender as principais propriedades físico-químicas dos ácidos nucleicos

Identificar as principais técnicas utilizadas em biologia molecular

Conhecer a metodologia de eletroforese de ácidos nucleicos

Conhecer o papel dessa técnica nas metodologias de Southern e Northern *blotting* e sua utilização

Compreender a clonagem gênica e a utilização de vetores e enzimas de restrição

Conhecer as bases teóricas e avaliar a importância da utilização da PCR e da PCR em tempo real

Compreender as bases teóricas e avaliar a importância da utilização da RT-PCR

Conhecer os diferentes métodos de sequenciamento de DNA

Conceitos-chave do capítulo

Absorbância
Ácido nucleico
Anelamento
Bioinformática
Biologia molecular
Conjugação
Desnaturação
Eletroforese
Enzima de restrição
Espectrofotometria
Genética molecular
Ionização
Northern *blotting*
Oligonucleotídios
Origem de replicação
PCR em tempo real
Plasmídios
Plasmídios R
Proteína
Reação em cadeia da DNA polimerase (PCR)
Sequenciamento de DNA
Sonda de DNA
Southern *blotting*
Temperatura de fusão ou Tm
Vetores de clonagem
Vetores de expressão

Introdução

A descoberta da estrutura do DNA, em 1953, pela dupla de pesquisadores Watson e Crick, levou a uma revolução científica admirável! Claro que esses pesquisadores não trabalharam sozinhos, e claro que isso foi, como toda descoberta científica, uma evolução de fatos, observações e ideias que culminou em um modelo comprovadamente correto. Como descrito em capítulos precedentes, a estrutura do DNA, a molécula da vida, é uma dupla hélice helicoidal, com duas cadeias complementares de orientações opostas constituídas por 3 componentes químicos: açúcar, fosfato e bases nitrogenadas. Aparentemente simples agora, essa descrição resultou em uma revolução biotecnológica, criando um ramo totalmente novo na ciência – a biologia molecular. A genética foi a ciência que mais se desenvolveu com o advento da biologia molecular.

Biologia molecular
Ramo da biologia que estuda a estrutura e a função do material genético, utilizando técnicas em química, bioquímica e microbiologia para a investigação de processos como a replicação celular e a expressão gênica

Denominamos genética molecular o ramo da genética dedicado aos mecanismos moleculares da hereditariedade, assim como às origens das patologias humanas.

Técnicas sofisticadas ou não têm sido amplamente utilizadas com o objetivo não apenas de aumentar o conhecimento nessas áreas, mas também, como não poderia deixar de ser, de obter lucro com o patenteamento de técnicas e moléculas. Mas isso é uma outra história… Neste capítulo, apresentaremos algumas técnicas mais usuais em genética molecular, assim como algumas técnicas mais avançadas e aplicações.

Genética molecular
Ramo da genética que estuda as bases moleculares da hereditariedade e das patologias humanas de origem genética

Técnicas em biologia molecular

Propriedades físico-químicas dos ácidos nucleicos

As tecnologias moleculares em biologia surgiram da integração de técnicas em química, bioquímica, microbiologia e genética. A convergência dessas disciplinas tornou possível a

Ácido nucleico
Macromolécula formada por subunidades denominadas nucleotídios, que contem a informação genética. Existem 2 tipos de ácidos nucleicos: DNA e RNA

Proteína
Molécula constituída por uma ou mais cadeias de aminoácidos, com uma estrutura tridimensional específica, que lhe confere determinada função biológica

Desnaturação
Processo pelo qual moléculas biológicas perdem sua estrutura quando submetidas a condições diferentes das que estavam no modo funcional. As condições alteradas podem ser temperatura, pH ou pressão

determinação da estrutura e das propriedades físico-químicas das principais moléculas envolvidas, tornando possível a criação de ferramentas de estudo importantes. Assim, os ácidos nucleicos e as proteínas formam as classes de moléculas biológicas sobre as quais os pesquisadores se debruçaram e tiraram, com brilhantismo, as metodologias correntemente utilizadas nos dias de hoje.

A estrutura dos ácidos nucleicos, o DNA e o RNA, já foi apresentada em capítulos anteriores (Capítulos 1 e 2). Citamos 3 propriedades físico-químicas dessa classe de moléculas, que são de fundamental importância para o desenvolvimento das técnicas moleculares: a complementariedade das bases nitrogenadas, a carga elétrica e a absorção de luz UV. Essas duas últimas propriedades também são compartilhadas pelas proteínas que, por sua vez, apresentam diversidade muito maior de formas e funções. Por ser um assunto muito extenso, não entraremos em detalhes sobre o estudo de proteínas. A Figura 10.1 mostra uma molécula de DNA cujas bases nitrogenadas se encontram em seu interior, associadas por complementariedade e um esqueleto exterior que é formado por resíduos de fosfato e pentose.

A complementariedade de bases do DNA dupla fita é formada pela união de bases pirimidinas e purinas dos nucleotídios, por meio de pontes de hidrogênio, formando os pares C-G e A-T ou A-U. Essa característica torna possível que a dupla fita se abra ao sofrer a desnaturação, por exemplo, por aumento de temperatura.

A desnaturação do DNA é o processo pelo qual um agente químico ou físico promove o rompimento das pontes de hidrogênio, levando à separação das fitas.

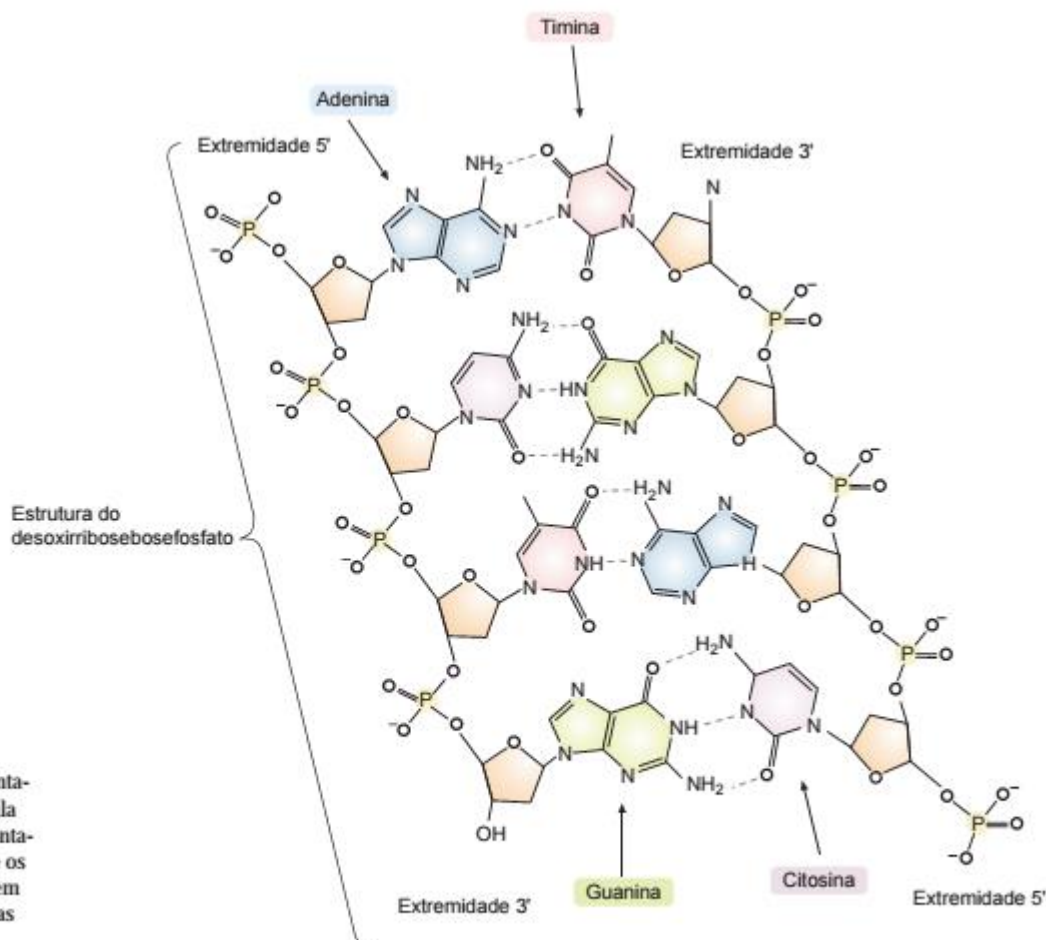



Figura 10.1 Esquema representativo da estrutura da molécula de DNA, no qual estão apontadas as bases nitrogenadas e os resíduos de fosfato. Estão em diferentes cores as estruturas em anel das bases.

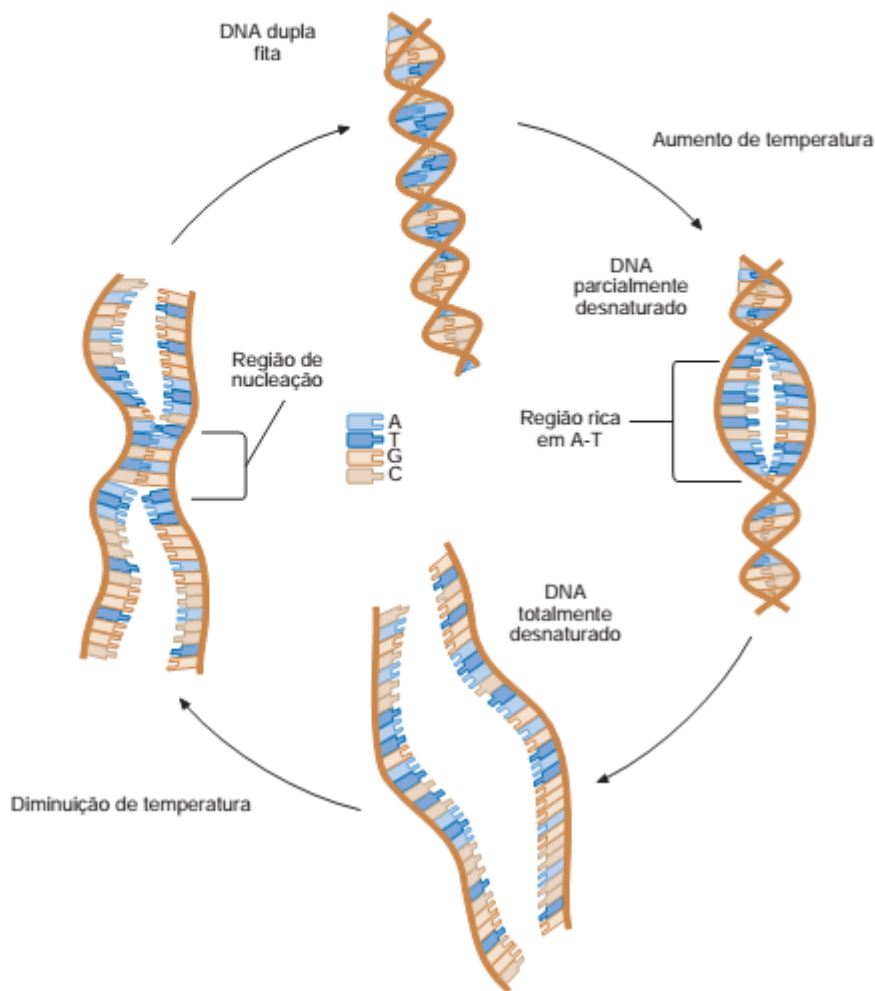



Figura 10.2 Representação da desnaturação e renaturação de uma molécula de DNA em razão da temperatura. Com o aumento da temperatura, as regiões ricas em pares AT se abrem primeiro e as ricas em GC, posteriormente. Com a diminuição da temperatura, as fitas tendem a se associar por complementariedade.

Cabe aqui ressaltar que as bases nitrogenadas podem promover a reassociação de duas fitas de ácidos nucleicos desde que estas sejam complementares entre si, ou seja, fitas de DNA de origens diferentes ou híbridos DNA-RNA podem ser formados com base nas sequências de nucleotídios complementares.

A Figura 10.2 mostra a desnaturação e a renaturação de uma molécula do DNA como função da temperatura.

Temperatura de fusão ou Tm
Temperatura na qual a metade das moléculas de DNA de uma amostra está sob a forma desnaturada

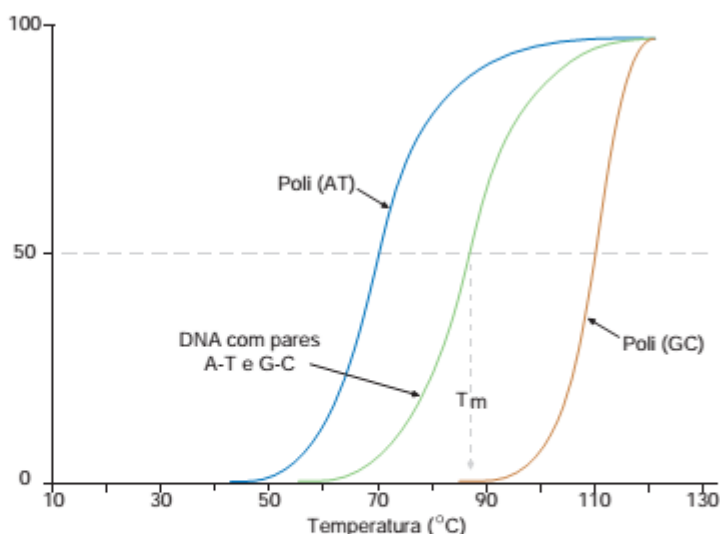
A temperatura de fusão ou Tm da molécula de DNA depende de sua constituição em nucleotídios, portanto, é característica de cada sequência de DNA. Essa temperatura indica o ponto em que uma amostra contendo moléculas de DNA apresenta 50% dessas desnaturadas.

Assim, a Tm é a temperatura na qual metade das moléculas de DNA está desnaturada e depende exclusivamente da sequência de nucleotídios.

Na Figura 10.3 podemos observar uma curva de desnaturação na qual moléculas de DNA formadas por diferentes sequências de nucleotídios se comportam diferentemente. Assim, uma sequência de DNA rica em G-C terá uma Tm mais alta que uma sequência rica em A-T. Note que em torno de 100ºC todas as moléculas se encontram em simples fita.

Ionização
Processo químico no qual são formadas estruturas químicas eletricamente carregadas pela perda ou pelo ganho de elétrons

Já a carga elétrica presente nos ácidos nucleicos é resultante da ionização dos resíduos de fosfato, que são responsáveis pela carga negativa característica dessas moléculas em solução. Cada nucleotídio da fita de ácido nucleico tem um resíduo de fosfato e, portanto, contribui com uma carga negativa. Finalmente, as bases nitrogenadas têm estrutura em forma de anel que podem, caracteristicamente, absorver a luz ultravioleta (UV) em comprimento de onda igual a 260 nm. Essa propriedade tem sido utilizada na quantificação dessas moléculas por

Figura 10.3 Curva de desnaturação de moléculas de DNA com conteúdo de nucleotídios diferentes.

Espectrofotometria
Método óptico de medida de absorção ou transmissão de luz do espectro visível ou invisível. Esse método é muito utilizado para identificação e quantificação de compostos

Absorbância
Quantidade de intensidade de luz que uma amostra absorve. A absorbância de uma solução é função linear da concentração de um composto

espectrofotometria. Ainda, de modo interessante, as proteínas absorvem a luz UV prioritariamente em outro comprimento de onda, ou seja, em 280 nm. Isso se deve à presença de anéis aromáticos em alguns aminoácidos (tirosina, triptofano e fenilalanina) e tem sido muito utilizado para mostrar a presença de proteínas contaminantes em uma solução de ácidos nucleicos, pela determinação da absorbância nos 2 comprimentos de onda.

A Figura 10.4 mostra o espectro ou curva de absorbância de uma solução de DNA e outra de proteína.

Desse modo, a desnaturação e a renaturação do DNA, assim como a carga elétrica e a absorção UV são propriedades dos ácidos nucleicos que são exploradas no desenvolvimento das tecnologias moleculares.

A seguir, apresentamos as técnicas mais importantes e usuais em biologia molecular.

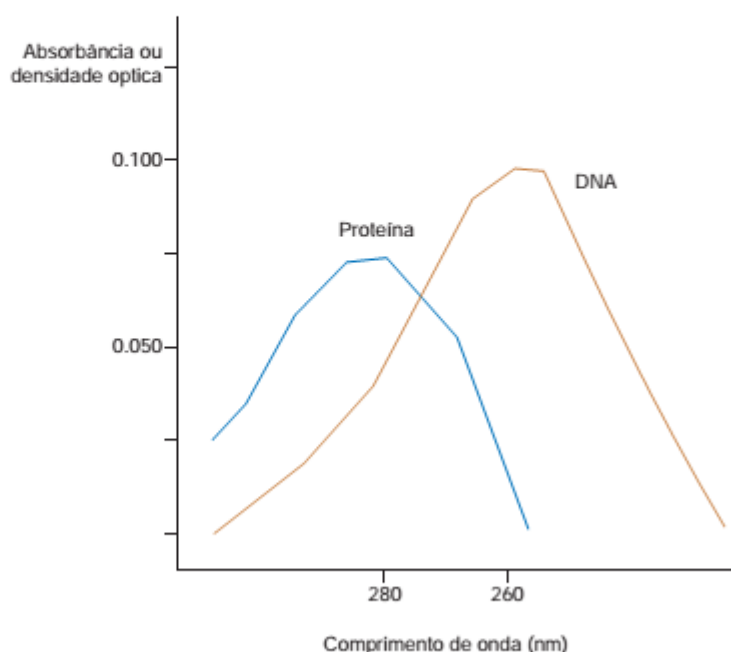



Figura 10.4 Curva da densidade óptica de uma solução de DNA e de proteína em relação ao comprimento de onda. O DNA absorve mais em 260 nm, ao passo que as proteínas em 280 nm.

Eletroforese de ácidos nucleicos em gel e o método de Southern *blotting*

Eletroforese
Método bioquímico que tem como principal objetivo a separação analítica ou preparativa de biomoléculas com base nas propriedades elétricas e de massa destas

Southern *blotting*
Técnica molecular utilizada na identificação de fragmentos de DNA específicos em uma mistura por meio da separação desses fragmentos em eletroforese em gel e transferência para membrana, seguida de hibridização com sondas de DNA marcadas, em fita simples e específicas

Sonda de DNA
Molécula de DNA de fita simples e marcada, utilizada para a identificação específica de fragmentos-alvo de DNA ou RNA, complementares

A eletroforese é um método muito utilizado em Bioquímica, que tem como principal objetivo a separação analítica ou preparativa de biomoléculas, com base nas propriedades elétricas e de massa destas. A eletroforese em gel é realizada em um suporte de matriz gelatinosa que pode ser do tipo gelatina, como a agarose, ou polímero, como a poliacrilamida. Esse suporte possibilita a passagem de íons e o deslocamento de moléculas carregadas, quando este é submetido a um campo elétrico, formado por um polo positivo e um polo negativo.

Assim, como já citado, cada nucleotídio contribui para uma carga negativa, fazendo com que o tamanho de uma molécula de DNA seja relativo ao número de cargas negativas.

A migração de uma molécula de DNA dupla fita na eletroforese é uma relação do seu tamanho, sendo que moléculas menores migram mais rapidamente que moléculas maiores.

O tamanho de uma molécula de DNA é uma medida do número de pares de bases que o forma. Na Figura 10.5 podemos observar como uma mistura de fragmentos de DNA se comporta quando submetidos à eletroforese em agarose. Cabe chamar aqui a atenção de que, para que os fragmentos de DNA sejam visualizados, o gel deve ser submetido à coloração, em geral com corantes como o brometo de etídio, que é um composto que se intercala na estrutura do DNA e se torna fluorescente em presença de luz UV.

Como vimos, uma mistura de moléculas clivadas ou de fragmentos de DNA pode ser separada por tamanho molecular, em uma eletroforese. No entanto, o que podemos fazer para identificar um fragmento específico, já que este vai migrar com inúmeras moléculas do mesmo tamanho, constituídas por diferentes sequências nucleotídicas?

A determinação de um fragmento de DNA que contenha uma sequência específica pode ser feita pelo método chamado de Southern *blotting*. Em 1975, o pesquisador E. M. Southern descreveu uma técnica baseada na transferência de moléculas de DNA, separadas anteriormente por eletroforese em gel de agarose, para uma membrana, no caso, de nitrocelulose. A transferência ocorre essencialmente por capilaridade, do gel de agarose para a membrana.

Na membrana, as moléculas de DNA se encontram desnaturadas e imobilizadas de modo a serem identificadas por uma sonda de DNA marcada, por exemplo, com isótopo radioativo.

A Figura 10.6 mostra a montagem do aparato de transferência do DNA do gel de agarose para a membrana.

A sonda de DNA nada mais é do que um fragmento de DNA com a sequência específica que se deseja identificar. Tal sonda, desnaturada, se ligará por complementariedade, especificamente ao fragmento relacionado estruturalmente.

A membrana contendo a sonda marcada associada ao fragmento específico a ser localizado será visualizada pela exposição em filme radiográfico. Na Figura 10.7, podemos observar a

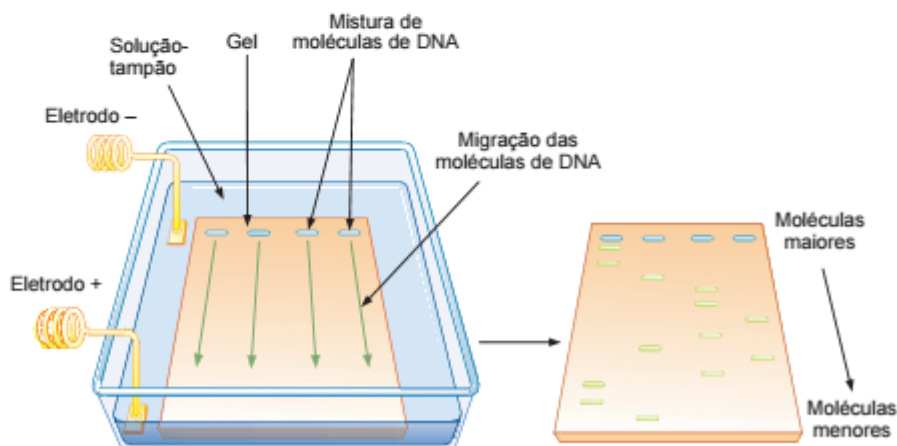
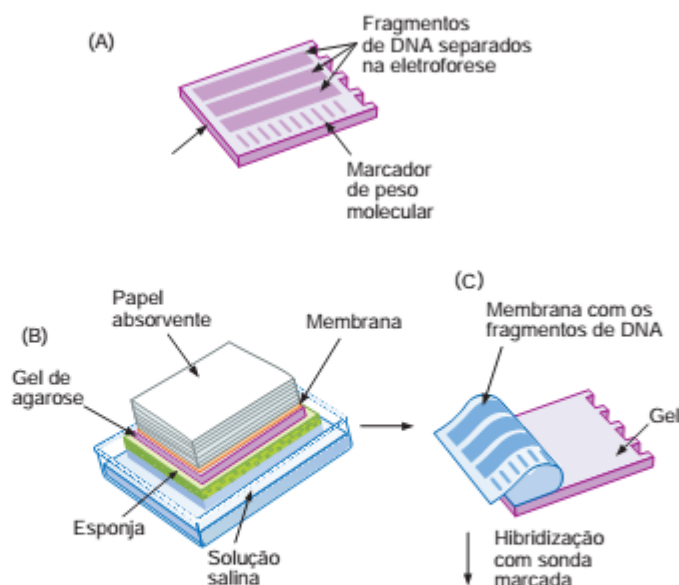



Figura 10.5 Representação esquemática do método de separação de uma mistura de moléculas de DNA por eletroforese.

Figura 10.6 Esquema representando o aparato de transferência do DNA para a membrana de nitrocelulose. (A) gel de agarose após a migração dos fragmentos de DNA; (B) aparato de transferência de DNA por capilaridade e seus componentes; e (C) membrana contendo os fragmentos de DNA após a transferência.

Enzima de restrição
Enzima de origem bacteriana, também denominada endonuclease de restrição, que age clivando o DNA dupla fita em sequências de bases específicas que são os sítios de restrição

Oligonucleotídios
Fragmentos curtos de DNA ou RNA, com até algumas dezenas de nucleotídios e que são, em geral, produzidos por síntese química, sob a forma de fita simples

representação de 2 exemplos em que uma sonda marcada se liga ao DNA-alvo, na técnica de Southern *blotting* discutida e na técnica conhecida como hibridização *in situ*.

Assim, é possível identificar, por exemplo, um fragmento de DNA que contenha um gene o qual se deseja estudar. Devemos chamar a atenção para o fato de que é absolutamente necessário que se obtenha uma sonda específica, que pode ser pela digestão de DNA por enzimas de restrição (ver a seguir) e purificação do fragmento de DNA ou pela construção de oligonucleotídios. Neste último caso, a sequência que se quer procurar tem de ser conhecida, pelo menos parcialmente.

Clonagem do DNA | Plasmídios e enzimas de restrição

A clonagem gênica é uma das técnicas básicas da chamada tecnologia do DNA recombinante que é especialmente importante ao se tratar de biotecnologia. De modo geral, a clonagem gênica

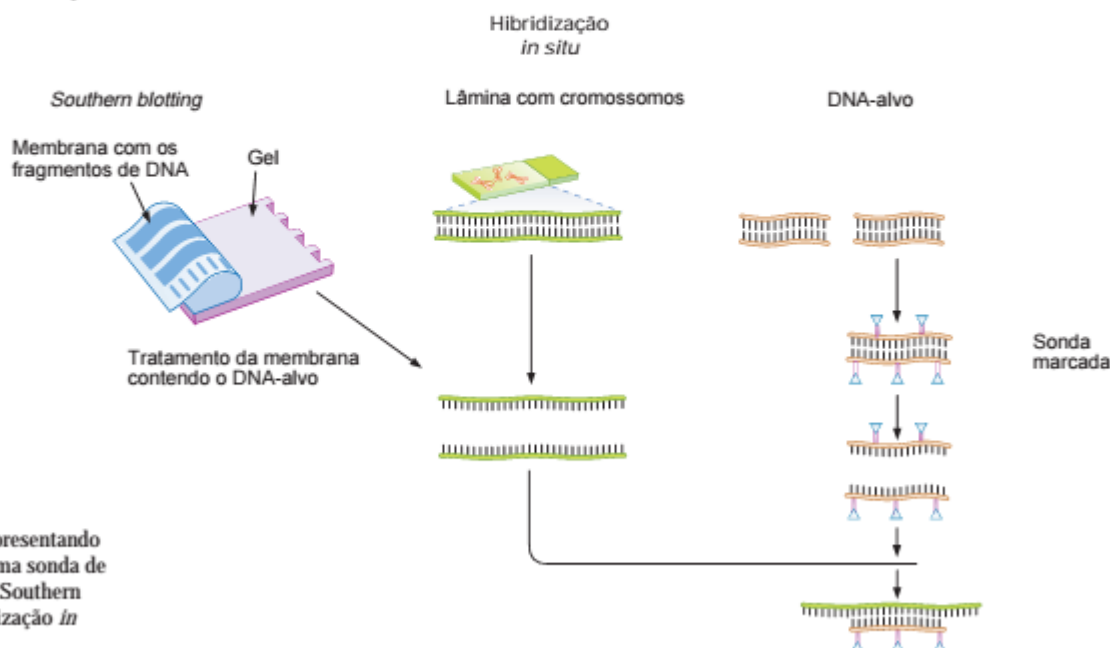



Figura 10.7 Figura representando a hibridização de uma sonda de DNA na técnica de Southern *blotting* e da hibridização *in situ*.

Plasmídios
Moléculas de DNA dupla-fita e circulares, extracromossômicos, presentes em alguns tipos de bactérias e que têm capacidade replicativa

Enzimas modificadoras de DNA
Enzimas com atividades ligadas à manipulação de DNA, de maneira a modificar a molécula de DNA *in vitro*, ao cortar, ligar, encurtar, alongar, copiar em DNA ou RNA, adicionar e remover grupos químicos

Plasmídios R
Plasmídios de origem bacteriana que apresentam genes que conferem resistência a alguns antibióticos

Conjugação
Processo pelo qual algumas bactérias podem trocar material genético, principalmente plasmídios

Origem de replicação
Sequência específica de nucleotídios presente no DNA genômico dos organismos, que marca o início da replicação do DNA

surgiu de trabalhos relevantes desenvolvidos no campo da microbiologia, principalmente o estudo e o desenvolvimento de plasmídios e a descoberta de enzimas modificadoras de DNA, entre elas as chamadas enzimas de restrição.

Os plasmídios são estruturas de DNA dupla-fita, presentes naturalmente em algumas bactérias e leveduras, e que carregam genes diferentes daqueles encontrados no genoma dos organismos hospedeiros.

Por exemplo, um grande problema em saúde pública tem sido o controle de resistência a múltiplos antibióticos que algumas bactérias patogênicas apresentam e que são transmitidos de bactérias resistentes a bactérias não resistentes. A resistência a antibióticos pode ser mediada pela presença de mutações gênicas que podem fornecer às bactérias portadoras a capacidade de sobreviverem em presença de antibióticos. Em alguns casos, as bactérias apresentam plasmídios com genes que, especialmente, produzem enzimas com propriedade de destruir as moléculas dos antibióticos, impedindo, assim, o seu papel. Esses plasmídios são denominados plasmídios R e podem ser transferidos para outras bactérias não resistentes por conjugação.

Assim, essas propriedades descritas fizeram dos plasmídios uma excelente ferramenta em biologia molecular, e pesquisadores passaram a utilizar essas estruturas para seleção de bactérias e para transferência de genes. Eles desenvolveram técnicas e selecionaram bactérias especialmente preparadas para receber essas estruturas. A bactéria mais utilizada nesses processos é da espécie *Escherichia coli*, que passou a ser a grande estrela em biotecnologia e que contém diferentes alterações em seu genoma, as quais as tornam funcionais para receber e transmitir plasmídios de interesse. Outra importante propriedade dos plasmídios é sua capacidade de autorreplicação. Todos os plasmídios têm uma origem de replicação e, portanto, podem se replicar dentro das bactérias, independentemente da replicação do genoma bacteriano. Você poderá encontrar mais detalhes sobre as origens de replicação no Capítulo 1, no item Replicação do DNA.

As enzimas de restrição, ou endonucleases de restrição, também existem naturalmente nas bactérias e têm como função primordial destruir o material genético estranho de, por exemplo, organismos do tipo *bacteriófagos* que, por acaso, penetrem nas bactérias.

As enzimas de restrição cortam o DNA dupla fita em locais específicos, chamados de sítios de restrição, clivando o esqueleto de açúcar-fosfato da cadeia de DNA, formando extremidades livres. Essas extremidades dos fragmentos de DNA formados podem ser de 2 tipos: extremidades coesivas e extremidades cegas.

As extremidades coesivas facilitam a ligação entre duas moléculas de DNA, já que podem se ligar por complementariedade. A Figura 10.8 mostra o exemplo de uma enzima de restrição, a EcoRI, que cliva a molécula de DNA e forma uma extremidade coesiva.

A denominação das enzimas de restrição é baseada no nome das bactérias de origem, assim, em nosso exemplo EcoRI é originária da *E. coli*. O estudo e a descoberta de inúmeras dessas

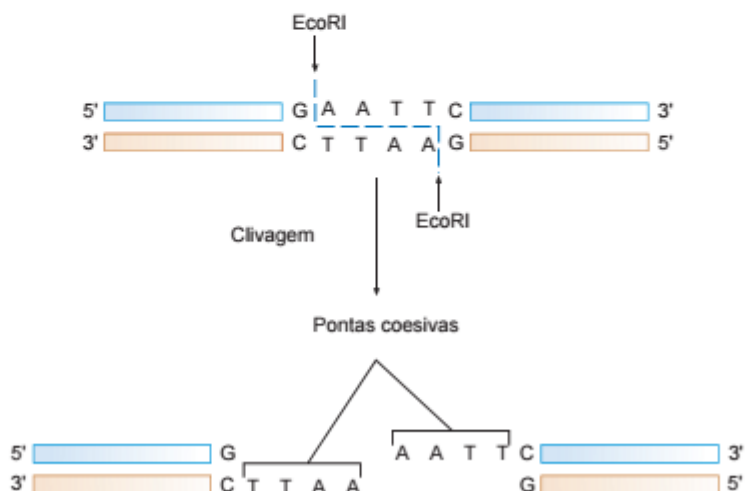



Figura 10.8 Esquema da clivagem do DNA pela enzima de restrição Eco RI no sítio de restrição, formando pontas coesivas.

enzimas e seus diferentes sítios de restrição foram fundamentais para o desenvolvimento de técnicas de clonagem. A Tabela 10.1 apresenta algumas dessas enzimas e sítios de clivagem específicos.

■ **Tabela 10.1** Principais enzimas de restrição, nome das bactérias que as originaram, sequência específica de reconhecimento e tipo de extremidade produzida.

Enzima	Origem	Sequência de reconhecimento	Tipo de extremidade
Ava I	*Anabaena variabilis*	C ▼ (C/T) C G (A/G) G	Coesiva
Bam HI	*Bacillus amyloliquefaciens*	G ▼ G A T C C	Coesiva
Bgl II	*Bacillus globigii*	A ▼ G A T C T	Coesiva
Eco RI	*Escherichia coli RY 13*	G ▼ A A T T C	Coesiva
Hae III	*Haemophilus aegyptius*	G G ▼ C C	Cega
hha I	*Haemophilus haemolyticus*	G C G ▼ C	Coesiva
Hind III	*Haemophilus influenzae Rd*	A ▼ A G C T T	Coesiva
Hpa I	*Haemophilus parainfluenzae*	G T T ▼ A A C	Cega
Kpn I	*Klebsiella pneumoniae*	G G T A C ▼ C	Coesiva
Pst I	*Providencia stuartii*	C T G C A ▼ G	Coesiva
Sma I	*Serratia marcescens*	C C C ▼ G G G	Cega
SstI	*Streptomyces stanford*	G A G C T ▼ C	Coesiva
Sal I	*Streptomyces albus G*	G ▼ T C G A C	Coesiva
Taq I	*Thermophilus aquaticus*	T ▼ C G A	Coesiva
Xma I	*Xanthamonas malvacearum*	C ▼ C C G G G	Coesiva

As enzimas de restrição, portanto, tornaram a manipulação do DNA possível e bastante simples. Assim, em um tubo de ensaio, moléculas de DNA podem ser cortadas especificamente, fornecendo fragmentos de DNA específicos, os fragmentos de restrição, que podem ser utilizados posteriormente, uma vez purificados, como sondas específicas ou, ainda, ser inseridos em plasmídios. Para que sejam inseridos em plasmídios, basta que estes sejam digeridos com a mesma enzima de restrição e que esta corte em apenas um local do plasmídio. Uma vez que o fragmento tenha o mesmo sítio, ou que seja compatível, este poderá ser inserido. Para completarmos o nosso novo plasmídio falta apenas a ação de outra enzima, a DNA ligase, cujo papel é catalisar a formação de ligação fosfato entre nucleotídios nas extremidades das moléculas de DNA. Assim, são construídas novas moléculas de plasmídios, modificados, com fragmentos de DNA que podem ter origens diferentes. A Figura 10.9 mostra a clonagem de genes em plasmídios bacterianos, sua introdução e a replicação em bactérias.

Agora, uma vez que você já entendeu o papel dos plasmídios modificados por pesquisadores, podemos denominar esses plasmídios de vetores.

Os vetores, portanto, são plasmídios modificados em sua estrutura para funcionar, de modo geral, na produção de múltiplas cópias de fragmento de DNA ou produzir um produto proteico, pela inserção de um gene de interesse.

Essas moléculas são fabricadas, segundo o interesse, em vetores de clonagem e vetores de expressão e podem ser encontrados para venda por empresas especializadas ou produzidos em laboratórios de pesquisa. A Figura 10.10 nos apresenta um plasmídio muito utilizado em técnicas de Biologia Molecular para clonagem e amplificação de fragmentos de DNA, denominado pUC19.

Vetores de clonagem
Plasmídios contendo em sua estrutura um fragmento de DNA de origem diferente e que foi inserido com o objetivo de ser introduzido e replicado em bactéria hospedeira

Vetores de expressão
Plasmídios que têm, além das estruturas de um vetor de clonagem, uma região promotora que irá dirigir a expressão de gene inserido

Amplificação de DNA *in vitro* | A reação em cadeia da DNA polimerase e a quantificação de seus produtos

Reação em cadeia da DNA polimerase (PCR)
Técnica para amplificação específica de DNA *in vitro*, pelo uso de iniciadores ou *primers* específicos que delimitam o segmento a ser copiado

A reação em cadeia da DNA polimerase, mais conhecida pela sigla PCR, foi desenvolvida nos anos de 1980 e revolucionou diversas áreas da biologia e da medicina. A PCR é utilizada para se obter a amplificação seletiva de determinada região de uma molécula de DNA.

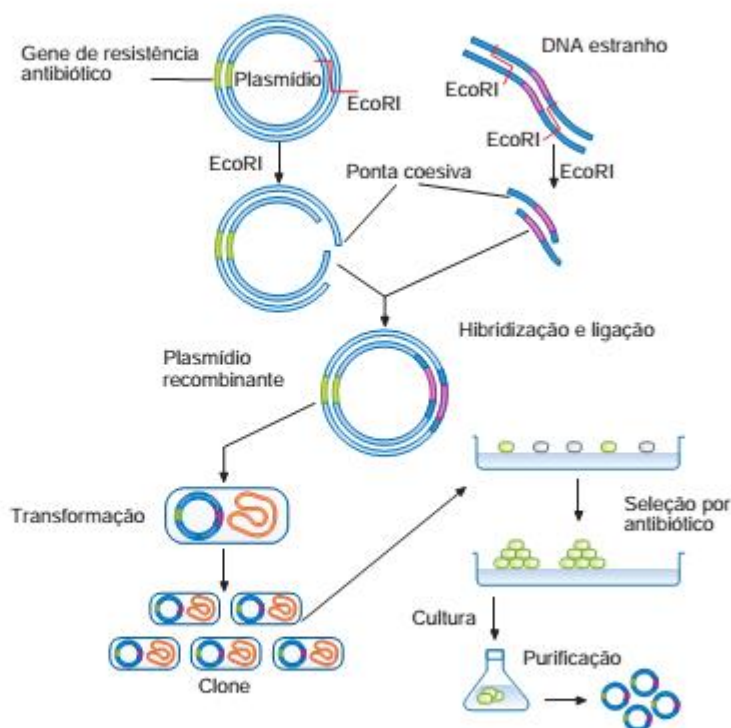



Figura 10.9 Etapas necessárias para que a clonagem em vetores bacterianos seja realizada.

Essa é uma técnica poderosa na qual, teoricamente, apenas uma única molécula de DNA pode servir de molde para amplificação, produzindo milhares de cópias da molécula-alvo. O idealizador dessa técnica foi o pesquisador americano K. Mullis, que ganhou o Prêmio Nobel de Química de 1993. As aplicações da PCR são inúmeras, podendo ser utilizadas para rastreamento e clonagem de genes específicos, diagnóstico de patógenos, teste de paternidade etc.

A PCR está baseada em 3 etapas de temperaturas diferentes que se repetem várias vezes, constituindo ciclos para a ação da enzima DNA polimerase.

Anelamento
Ligação de oligonucleotídios iniciadores ou *primers* ao DNA-alvo, por complementariedade

Assim, as etapas são: desnaturação térmica do DNA-molde, em geral 94°C; hibridização **ou** anelamento *dos oligonucleotídios ou iniciadores*, **também conhecidos como** ***primers*****, a cada** uma das fitas do DNA-molde; a temperatura dessa etapa é próxima à temperatura de fusão ou Tm (ver anteriormente) e depende do par de iniciadores utilizados; e polimerização das fitas **novas de DNA, a partir de cada um dos iniciadores.**

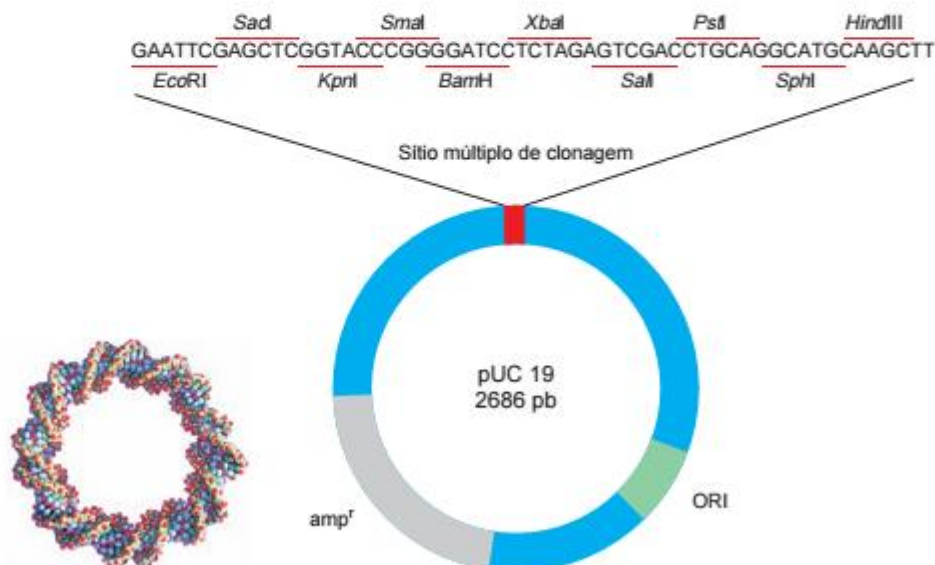



Figura 10.10 Representação do plasmídio pUC19 e suas principais características. No detalhe, uma representação do pUC19 em dupla fita.

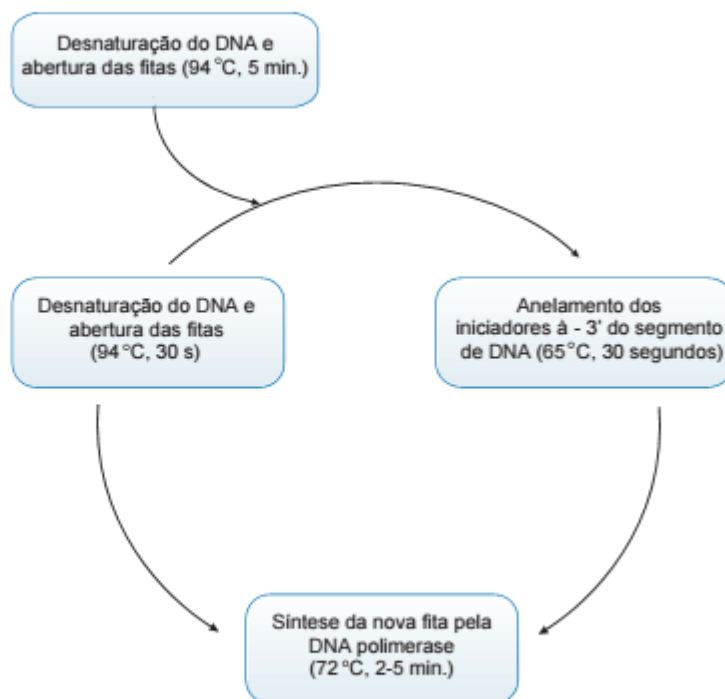



Figura 10.11 Esquema que representa as diferentes etapas sucessivas da PCR.

A Figura 10.11 mostra como essas etapas se sucedem.

A enzima de escolha é a DNA polimerase da bactéria *Thermus aquaticus*, ou taq polimerase, que apresenta termoestabilidade. A descoberta de bactérias que habitam fontes termais forneceu a essa metodologia um grande avanço pois, inicialmente, a enzima utilizada era a DNA polimerase da *E. coli* que, a cada etapa de desnaturação, tinha de ser reposta, visto que ela era destruída. Isso tornava a técnica dispendiosa e demorada. Assim, as enzimas termoestáveis, associadas ao desenvolvimento de equipamentos denominados termocicladores, fizeram uma grande revolução metodológica ainda nos anos 1980. Agora, iremos comentar mais detalhadamente as etapas da PCR e, para mais esclarecimentos do processo de replicação do DNA, consulte o Capítulo 1.

As DNA polimerases usam como molde para a replicação do DNA uma molécula em fita simples. Assim, a primeira etapa para o início da PCR é a etapa de desnaturação completa do DNA. Conforme exposto no item sobre as propriedades dos ácidos nucleicos, as moléculas do DNA podem ser separadas pelo aumento da temperatura, sendo que, de maneira geral, uma temperatura em torno de 94°C, por 5 min, leva à desnaturação completa das moléculas de DNA em uma mistura dessas. No entanto, para que a DNA polimerase dê início à síntese da fita complementar, ela reconhece e se liga ao DNA dupla fita. Portanto, para que a síntese ocorra em um tubo de ensaio, é necessário que sejam construídos os iniciadores específicos que irão se anelar, marcando o trecho da molécula que será copiada.

A síntese de iniciadores é crítica para o sucesso da PCR. Em geral, esses iniciadores são oligonucleotídios contendo entre 18 e 24 nucleotídios, com temperatura de anelamento ou Tm semelhantes.

Existe uma fórmula para o cálculo aproximado do Tm de iniciadores. Essa fórmula leva em consideração o fato de que o par A-T precisa de menor temperatura para se desfazer, pois apresenta duas pontes de hidrogênio para sua ligação, que o par G-C, que tem 3. Assim, pode-se determinar o Tm por:

$$Tm = (4 \times [G + C]) + (2 \times [A + T])°C$$

em que [G + C] é o número dos nucleotídios G e C e [A + T] é o número de A e T presentes na sequência de DNA. Ainda, os iniciadores não podem anelar entre si nem internamente, formando estruturas em alça, do tipo *hairpin* nas temperaturas utilizadas, levando a uma reação

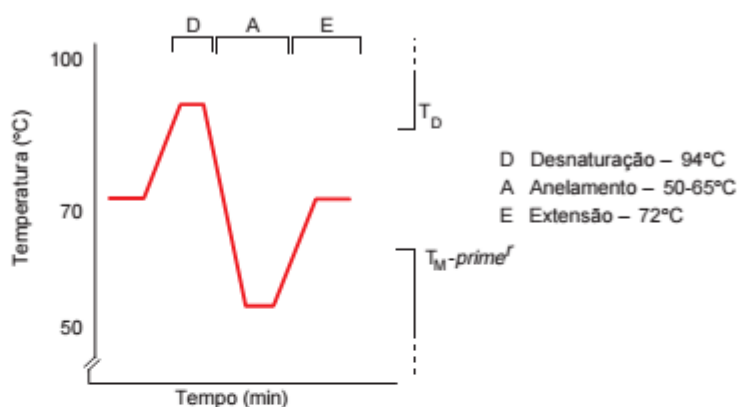



Figura 10.12 Gráfico de temperatura relacionado com as etapas da PCR.

improdutiva. Para evitar esses problemas e facilitar a utilização dos iniciadores, devem ser adotados programas de computador ou *softwares* especializados na construção de iniciadores adequados à região a ser amplificada. Uma vez anelados os iniciadores, segue-se a etapa de alongamento ou síntese da nova fita de DNA. Para que a PCR aconteça, devem existir no meio de reação os seguintes componentes: enzima taq DNA polimerase, desoxinucleotídios trifosfato (dATP, dCTP, dGTP e dTTP), iniciadores, íons Magnésio, solução-tampão da enzima e, ainda, a temperatura deve estar adequada para a atividade enzimática, sendo para a taq polimerase de 72°C. Seguindo-se à etapa de alongamento ou extensão, existe uma nova etapa de desnaturação a 94°C, reiniciando-se, assim, novo ciclo. A Figura 10.12 apresenta o gráfico das mudanças de temperatura durante um ciclo da PCR.

Assim, os ciclos de desnaturação-anelamento-extensão serão repetidos quantas vezes necessárias. Em condições ideais, a partir do terceiro ciclo, quando as moléculas longas não são mais amplificadas e o produto tem o tamanho desejado, este será amplificado em escala exponencial.

Assim, depois de 30 ciclos completos, haverá mais de 130 milhões de moléculas de DNA, com a sequência e tamanho desejados, com base em uma molécula-molde.

Desse modo, de quantidades muito pequenas de DNA-molde pode-se obter microgramas de produtos da PCR. A Figura 10.13 apresenta os produtos da PCR esquematicamente.

Apesar de a PCR ser uma técnica bastante simplificada, pesquisadores que a utilizam no trabalho se deparam diariamente com alguns problemas importantes. Entre eles podemos citar a especificidade, a introdução de bases erradas pela polimerase e a presença de contaminação por DNA exógeno. O primeiro problema diz respeito à especificidade que depende não só da presença de oligonucleotídios iniciadores adequados, mas também da utilização da temperatura de anelamento correta. A hibridização de moléculas de DNA complementares é dependente da temperatura, como já discutido. Uma temperatura mais alta impede o anelamento, tornando a PCR não produtiva, e uma temperatura mais baixa possibilita a formação de anelamento inadequado com híbridos mal pareados, que serão amplificados pela polimerase. Esse processo formará produtos não desejados, podendo estes serem avaliados por meio de eletroforese em agarose, por seu tamanho molecular (ver item anterior). Existem técnicas em biologia molecular que se utilizam do anelamento parcial de iniciadores. Na chamada PCR-RAPD, genomas de animais de espécies diferentes podem ser comparados, mostrando padrões diferentes de produtos de PCR. A introdução de nucleotídios errados na fita gerada na PCR pode acontecer com todas as enzimas DNA polimerases utilizadas comercialmente. Estima-se que a taxa de incorporação errada da Taq polimerase fique em torno de 1 para cada 2×10^4 nucleotídios. Como esse erro provocado é aleatório, este só irá interferir caso os produtos da PCR venham a ser clonados, o que levaria à possível seleção do produto alterado e sua posterior amplificação. Para minimizar esse problema, são introduzidas DNA polimerases com a propriedade de correção, pela atividade exonuclease de 3′ para 5′, no meio de reação fazendo com que a taxa de erro diminua. Um exemplo dessas enzimas é a Pfu, obtida comercialmente. Por fim, o problema da contaminação é um dos maiores desafios, por exemplo, em pesquisa forense.

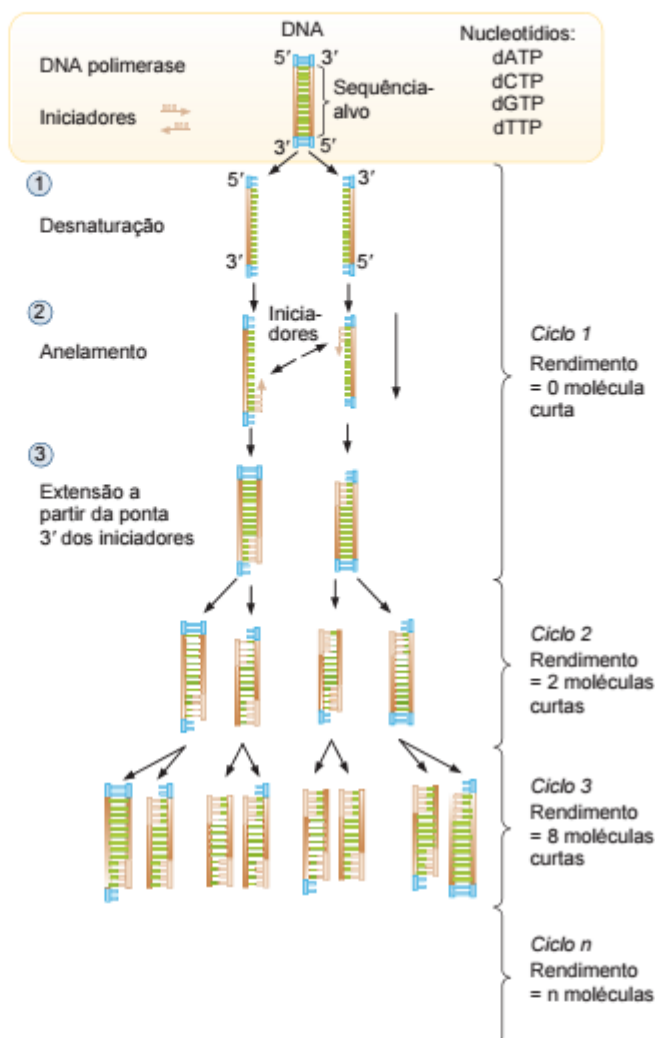



Figura 10.13 Representação da PCR desde o primeiro ciclo e a formação dos produtos.

PCR em tempo real
Tipo de PCR que se baseia na detecção dos produtos formados pela ação da polimerase durante o processo

Por trabalhar com quantidades ínfimas de DNA, na busca de provas para identificação de um indivíduo, que podem ser fios de cabelo, muitos ciclos devem ser aplicados, o que promove a amplificação de moléculas de DNA não desejado. Medidas importantes, como a utilização da técnica em ambiente livre de DNA e o uso de luvas, gorros e aventais utilizados apenas no local de manuseio dos reagentes da PCR, são de fundamental importância. Em geral, a introdução de mais um tubo na PCR, contendo na mistura reacional todos os reagentes menos o DNA-molde, ajuda na identificação de DNA contaminante, servindo como controle para a reação. Um importante aspecto decorrente da PCR é a possibilidade de quantificar o DNA-molde inicial. A PCR quantitativa ou qPCR foi desenvolvida baseando-se na detecção dos produtos da PCR, de maneira contínua, durante o desenvolvimento da reação de amplificação. Por isso, essa metodologia também é chamada de **PCR em tempo real**.

A quantificação dos produtos é feita por meio da fluorescência emitida por marcador associado a cada molécula sintetizada pelos sistemas de detecção. Assim, alguns métodos e equipamentos foram desenvolvidos no intuito de aperfeiçoar a quantificação. Os mesmos reagentes da PCR convencional são utilizados acrescidos de fluorocromos, que irão se intercalar na molécula de DNA recém-sintetizado. Isso pode ocorrer de duas maneiras, com corantes que se intercalam no DNA, como, por exemplo, no método SyberGreen, ou por meio de sondas específicas com fluorocromos, como na metodologia conhecida como Taqman. Os termocicladores são especiais, com um leitor de fluorescência emitida pela excitação dos fluorocromos por *lasers*, sendo essa emissão, portanto, detectada e proporcional à quantidade de produto formado.

Em linhas gerais, pode-se evidenciar 3 fases durante a PCR: a fase inicial, na qual os produtos não são formados, seguida pela fase exponencial, na qual pode-se observar a duplicação

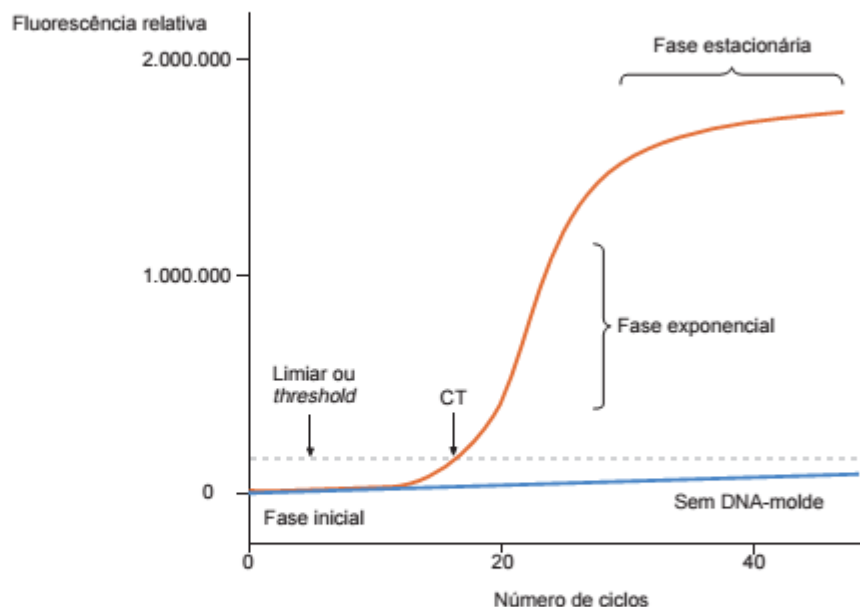



Figura 10.14 Gráfico representando uma curva típica da PCR em tempo real, mostrando as principais fases e índices da qPCR.

de moléculas-molde e a fase estacionária em que se observa uma parada na produção de novos produtos. Essa fase final se deve ao esgotamento de reagentes no meio da reação, não importando o número de ciclos realizados. Assim, se mais moléculas de DNA estiverem presentes no início da PCR, menos ciclos serão necessários para que a reação chegue à fase exponencial. A Figura 10.14 mostra uma curva da PCR, na qual estão evidenciadas as 3 fases descritas.

Na PCR em tempo real, determina-se, de maneira bastante acurada, a quantidade de DNA no início da PCR, por meio de comparação do número de ciclos.

A Figura 10.15 ilustra curvas feitas pela detecção de produtos formados e número de ciclos, partindo-se de uma diluição em série de uma solução de DNA.

Você já pode prever o grande avanço técnico alcançado pelo desenvolvimento da PCR em tempo real. Na clínica, pode-se obter com mais precisão, por exemplo, a carga viral de indivíduos infectados. Em pesquisa, essa metodologia ajudou na obtenção de resultados importantes, como, por exemplo, pela determinação do número de cópias de genes e na quantificação de RNA, como veremos a seguir.

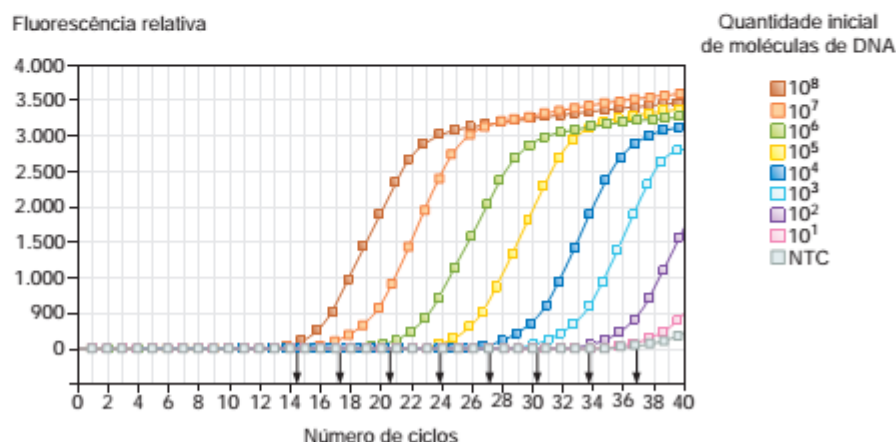



Figura 10.15 Gráfico mostrando o efeito da diluição em série da quantidade inicial de DNA na fluorescência produzida.

Análise da expressão gênica | Northern *blotting*, a reação da transcriptase reversa seguida da reação da polimerase e a tecnologia de microarranjos

O estudo da expressão gênica é uma etapa obrigatória na determinação da função de genes descobertos pelo estudo de genomas ou para a avaliação da modulação de genes nas células em diferentes estados. A expressão gênica foi objeto do Capítulo 2 no qual você poderá encontrar detalhes. De fato, esse é um termo utilizado para designar o resultado final da atividade gênica que expressa, literalmente, o que está determinado pelos vários níveis de controle da formação do produto final que pode ser uma proteína ou um RNA.

Northern *blotting*
Técnica de identificação de mRNA em uma mistura dessas moléculas por meio da separação do RNA total em eletroforese em gel e transferência para membrana, seguida de hibridização com sondas de DNA marcadas, em fitas simples e específicas

O primeiro método a ser abordado é a técnica conhecida como Northern *blotting*. Ela foi desenvolvida como consequência natural do Southern *blotting*, e seu nome é um jogo de palavras feito com o nome do pesquisador E.M. Southern, responsável pela descrição do método. O Northern *blotting* segue etapas semelhantes às descritas no item 2.2, sendo que, neste, o RNA total de um tecido ou de células em cultura é isolado, e moléculas de RNA são transferidas para membrana de náilon, sendo o transcrito identificado pelas sondas de DNA, marcadas e específicas. Cabe aqui chamar a atenção de que essa sonda de DNA deve ser correspondente ao DNA complementar do transcrito maduro que, em última análise, é o objeto de estudo dessa técnica.

A Figura 10.16 mostra o esquema das etapas utilizadas no Northern *blotting*.

Outro método muito utilizado no estudo da expressão gênica é o procedimento conhecido como reação da transcriptase reversa seguida da reação da polimerase (RT-PCR) que foi uma poderosa inovação técnica.

Na RT-PCR ocorre a amplificação por PCR do cDNA (DNA complementar ao mRNA) obtido pela ação de uma enzima denominada transcriptase reversa.

Também conhecida como DNA polimerase RNA-dependente, a transcriptase reversa (RT) foi identificada em retrovírus e tem como atividade a síntese de uma molécula de DNA a partir de uma molécula de RNA-molde. De modo semelhante às DNA polimerases, para o início da síntese do DNA complementar (cDNA) pela RT é necessária a presença de iniciadores ou *primers*. Estes, por sua vez, podem ser aleatórios ou direcionados. Os primeiros são, em geral, hexâmeros que vão se associar aleatoriamente aos RNAs e os direcionados são oligonucleotídios formados apenas de desoxitimidina, ou poli(dT), que vão se associar especificamente à cauda

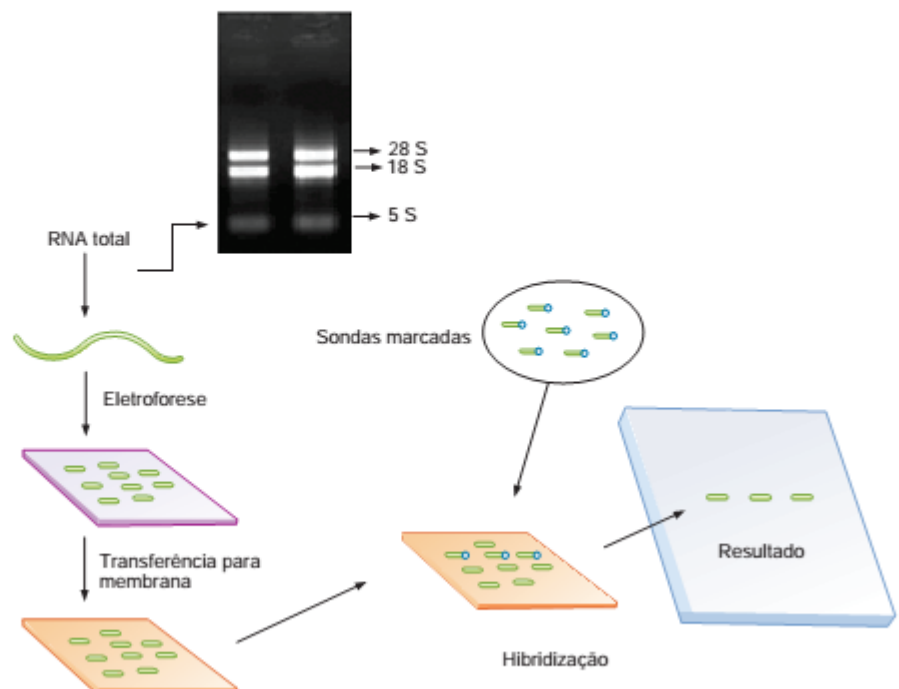



Figura 10.16 Etapas da metodologia de Northern *blotting*. No detalhe, um gel de agarose mostrando uma separação de RNA total típica.

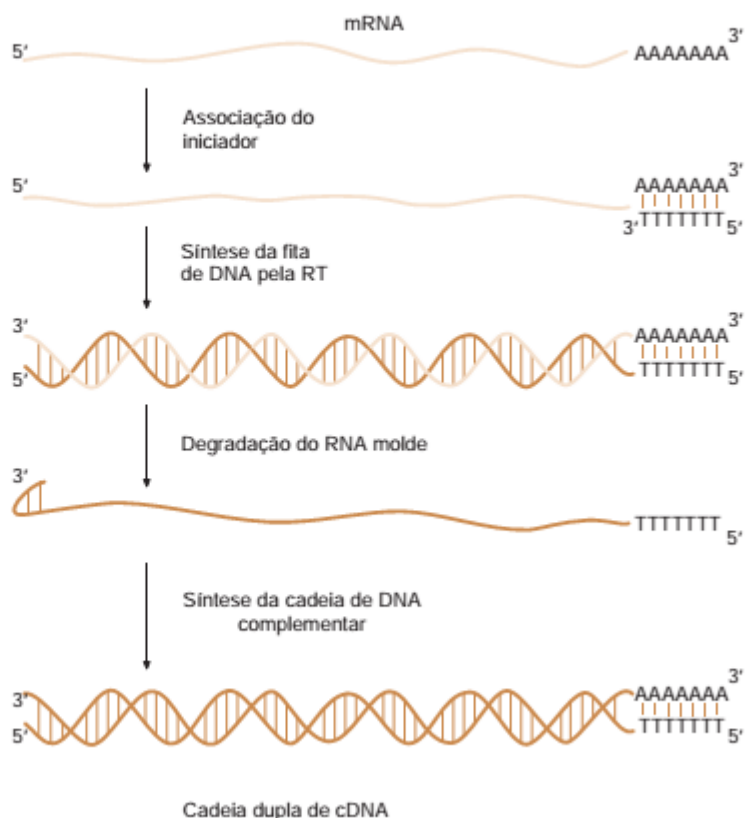



Figura 10.17 Etapas da transcriptase reversa seguida da reação da polimerase (RT-PCR).

poliA dos RNAs mensageiros maduros. Assim, com base em moléculas variadas de cDNA, pela PCR específica, pode-se amplificar apenas o cDNA desejado. A Figura 10.17 mostra as principais etapas da RT-PCR.

Além de relativa facilidade técnica, esse procedimento possibilita a detecção de transcritos com pequeno número de cópias e/ou de pequenas quantidades de células, por exemplo, no estudo da expressão gênica nos primeiros estágios do desenvolvimento embrionário. Em virtude de sua alta especificidade e sensibilidade, tem sido a técnica de escolha para o estudo de diversos produtos gênicos.

Para a análise quantitativa dos mRNA pesquisadores desenvolveram a RT-PCR quantitativa (qRT-PCR), na qual a PCR em tempo real está acoplada à reação da transcriptase reversa.

Desse modo, pela determinação da quantidade de transcritos produzidos, pode-se observar alterações na expressão gênica e compará-las. Nesse sentido, um grande avanço no estudo da expressão gênica foi obtido com o desenvolvimento da metodologia de microarranjos.

Os microarranjos nada mais são do que arranjos de moléculas de DNA predeterminados, usualmente oligonucleotídios de 20 a 25-mer, ligados quimicamente e colocados em linha em uma superfície que serve de base para hibridização e posterior detecção.

Essa superfície pode ser uma membrana de náilon ou uma lâmina de quartzo, também chamadas de *chips*, em referência aos *chips* de computadores. A amostra contendo a mistura de moléculas marcadas com radioisótopo ou fluorocromos é colocada nos microarranjos para que ocorra a hibridização de maneira específica. Após serem retiradas as moléculas que não se associaram, a detecção é feita por scanner a *laser* acoplado a um computador que identificará os oligonucleotídios, determinando, assim, o tipo de DNA ou cDNA presente na amostra. Essa técnica foi descrita no final dos anos 1980 e vem sendo utilizada para diversas finalidades, como no estudo da variabiliade genética. Em relação ao estudo da expressão gênica, essa técnica se mostrou extremamente útil para a detecção de expressão diferencial de células submetidas a

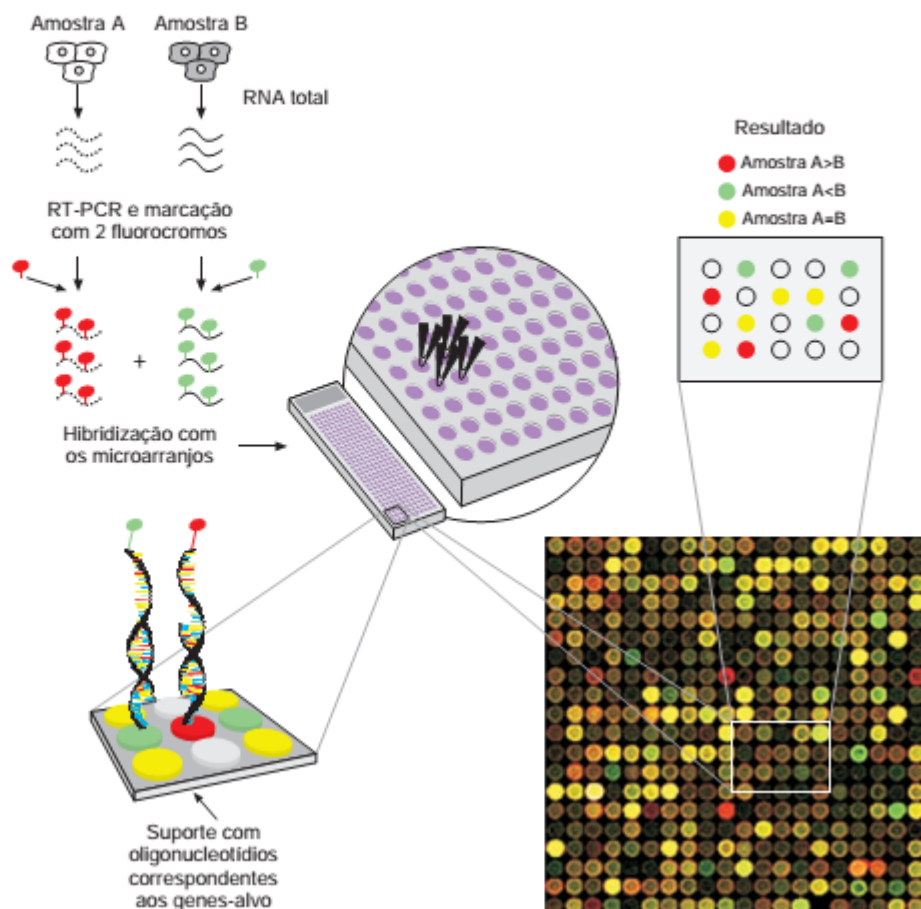



Figura 10.18 Metodologia de microarranjos no estudo comparativo da expressão gênica entre duas amostras distintas.

diferentes tratamentos ou de tecidos em estágios diferentes de desenvolvimento. A Figura 10.18 apresenta os detalhes da técnica de microarranjos no estudo da expressão gênica diferencial entre duas amostras diferentes.

Sequenciamento de DNA

O conhecimento da sequência correta de nucleotídios de um fragmento de DNA, de um gene ou mesmo do genoma de um organismo, se tornou etapa indispensável na pesquisa em genética desde a descoberta da estrutura do DNA em 1953. Com o desenvolvimento de técnicas cada vez mais sofisticadas, o sequenciamento de DNA vem se tornando mais acessível e, como consequência, um enorme acúmulo de dados sobre os genomas e sua variabilidade vem acontecendo. No entanto nem sempre foi assim.

Sequenciamento de DNA
Conjunto de técnicas com o objetivo de determinar a sequência de bases de nucleotídios de fragmentos de DNA

O primeiro método de sequenciamento rápido foi o *plus-minus*, altamente complexo e que foi utilizado em 1977 para o sequenciamento completo do primeiro organismo, o bacteriófago φX174. Outro método desenvolvido com sucesso foi o químico, conhecido pelo nome de seus autores, A. Maxam e W. Gilbert.

O método de Maxam-Gilbert consiste na modificação química e diferencial das purinas (A e G) e das pirimidinas (C e T) de moléculas de DNA marcadas radioativamente, seguida de clivagem, resultando na formação de fragmentos de DNA de diferentes tamanhos.

Os fragmentos são, então, separados por eletroforese e identificados por meio de autorradiografia. Esse método foi utilizado para a dedução da sequência completa do vírus SV40 e do plasmídio recombinante pBR322 que havia sido recentemente produzido por pesquisadores. Apesar do enorme sucesso, o método Maxam-Gilbert foi suplantado pelo método enzimático, desenvolvido pela equipe do Dr. Sanger, e descrito também em 1977 na revista Nature.

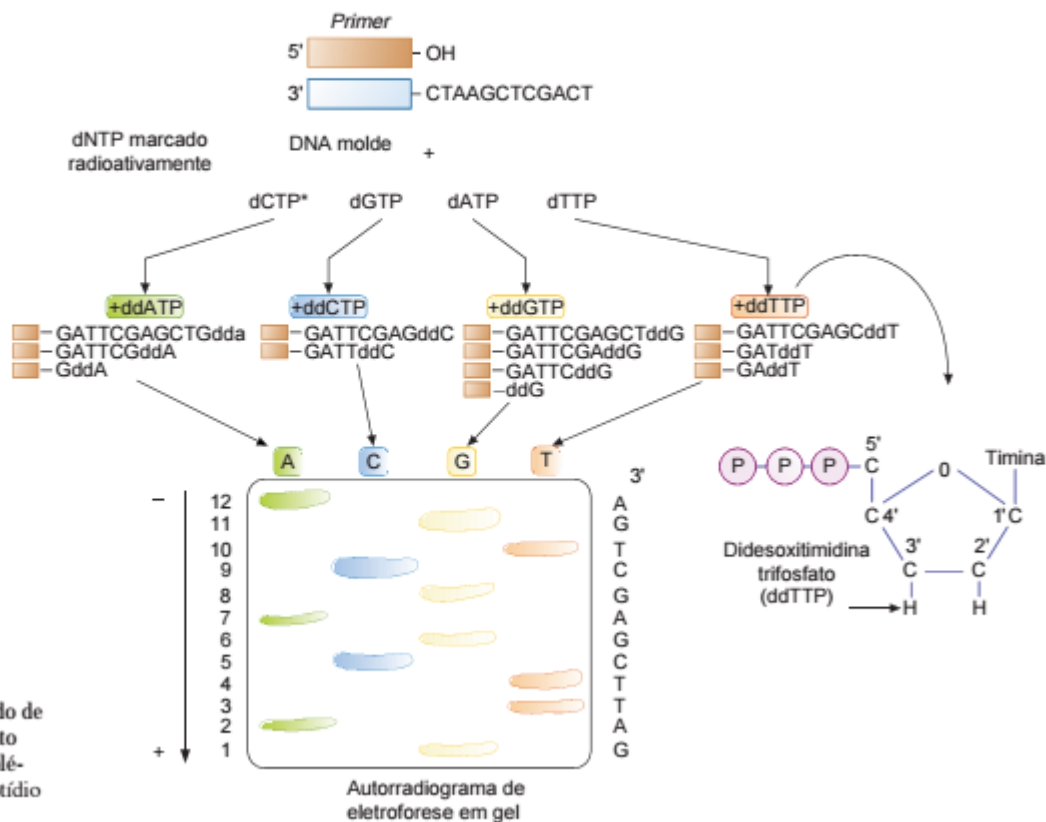



Figura 10.19 Etapas do método de Sanger para sequenciamento de DNA. No detalhe, a molécula de um didesoxinucleotídio (ddTTP).

O método de Sanger (enzimático ou de terminação de cadeia) baseia-se na síntese de uma fita de DNA a partir do DNA-molde a ser sequenciado, utilizando no meio reacional os didesoxinucleotídios – nucleotídios modificados que promovem a terminação prematura da nova cadeia.

Essas moléculas não têm a hidroxila na posição 3′ e podem ser incorporadas ao DNA nascente pela presença de um trifosfato livre na posição 5′; no entanto, não podem formar a ligação fosfodiéster com o próximo nucleotídio. Assim, são necessárias 4 reações diferentes, cada uma contendo todos os elementos para a síntese da nova fita de DNA e um didesoxinucleotídio diferente, quais sejam: ddCTP, ddGTP, ddTTP e ddATP. Os diferentes fragmentos formados são, então, separados por eletroforese de alta resolução. A Figura 10.19 mostra as etapas do sequenciamento pelo método de Sanger.

Uma importante melhoria feita a essa metodologia foi aplicada pelos pesquisadores M. Hunkapiller e L. Hood, com base na utilização de fluorocromos associados às moléculas de didesoxinucleotídio. Cada ddNTP tinha um fluorocromo que emitia uma cor diferente que, por sua vez, era detectada por equipamento a *laser*. Uma só mistura reacional era realizada para a obtenção da sequência. A Figura 10.20 apresenta um esquema no qual os fragmentos originados pela síntese da nova fita de DNA em presença de ddNTPs são separados e detectados, resultando no sequenciamento da cadeia-molde.

Muitos foram os avanços realizados pelos pesquisadores pela utilização da metodologia de sequenciamento em eletroforese capilar com base na técnica de Sanger para a elucidação da informação genética. Entretanto, tornou-se evidente a necessidade de se obter grande quantidade de dados, incluindo a determinação de genomas completos. Deu-se início, então, à corrida para o desenvolvimento de métodos mais rápidos, mais baratos e confiáveis. Assim, iniciou-se a era da genômica, cujo ápice científico foi a determinação da sequência do genoma humano em 2003, 50 anos depois da descoberta da estrutura do DNA. O sucesso alcançado para esse ambicioso projeto está relacionado ao desenvolvimento de novos vetores de clonagem, sequenciadores com maior desempenho e, seguindo de perto, a análise de dados por computadores cada vez mais potentes. Com o advento da bioinformática, esses elementos resultaram na

Bioinformática
Conjunto de metodologias associadas ao uso de computadores, programas e modelos matemáticos em processamento, integração e ordenação de dados biológicos

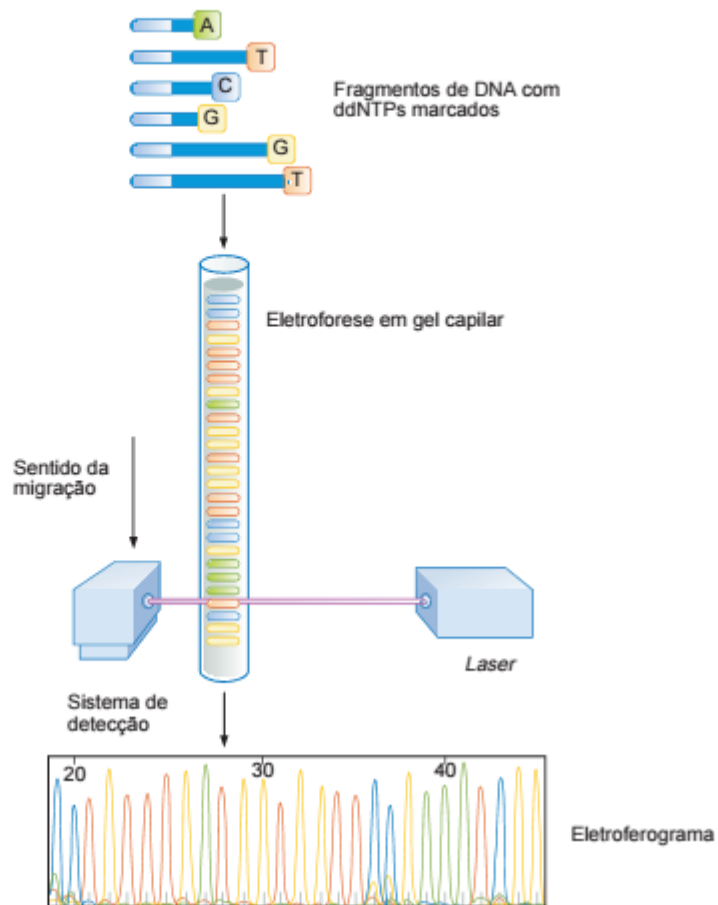



Figura 10.20 Representação de sequenciamento automático, no qual os fragmentos gerados pela síntese da nova fita possuem fluorocromos e são detectados por equipamento a *laser*.

descrição de genomas completos de organismos, assim como na comparação entre dados de sequências obtidos de centenas ou mesmo de milhares de segmentos de DNA.

Algumas novas técnicas, como o pirosequenciamento, comercializado pela Roche®, e o sequenciamento por síntese, da Illumina, levaram à evolução dos sequenciamentos para os estudos em larga escala. Esses métodos se baseiam no sequenciamento de milhares de segmentos de DNA ao mesmo tempo, o que promove a possibilidade de produzir enorme quantidade de sequências em menor tempo.

Esse tipo de abordagem levou ao desenvolvimento dos estudos chamados de análises em larga escala e à chamada tecnologia de sequenciamento de última geração (Next Generation Sequence Technology – NGS).

G·F Genética em Foco

Genética e bioinformática

A genética molecular vem avançando para os estudos conhecidos como de larga escala. Graças ao desenvolvimento das grandes redes computacionais, como a internet, estamos vivendo o momento de internacionalização e ciência sem fronteiras. O resultado mais imediato dessa nova abordagem é a formação de inúmeros consórcios de pesquisadores que, em diversos países e centros de pesquisa, podem reunir dados que são acumulados em grandes bancos. Exemplos desses consórcios são: o International Histocompatibility Working Group

(http://www.ihwg.org/); International HapMap Project (http://hapmap.ncbi.nlm.nih.gov/); Human Genome Diversity Project (http://hsblogs.stanford.edu/morrison/human-genome-diversity-project/). Podemos afirmar que, após o projeto Genoma Humano, vem ocorrendo o desenvolvimento de grandes bancos de dados em genoma, transcriptoma e proteoma que, por sua vez, são acumulados em maior velocidade do que os pesquisadores podem utilizá-los. Assim, surge a figura do bioinformata, ou seja, o profissional que está envolvido tanto no desenvolvimento e na utilização de programas de análise computacionais quanto nos temas e problemas em biologia. A bioinformática surge, então, da interação entre o conhecimento de engenharia de *softwares*, a matemática, a estatística, a ciência da computação e a biologia molecular.

RESUMO

- A genética molecular é o ramo da genética dedicado aos mecanismos moleculares da hereditariedade, assim como das origens genéticas das patologias humanas
- As tecnologias moleculares em biologia surgiram da integração de técnicas em química, bioquímica, microbiologia e genética
- A desnaturação do DNA é o processo pelo qual um agente químico ou físico promove o rompimento das pontes de hidrogênio, levando à separação das fitas
- A Tm é a temperatura na qual metade das moléculas de DNA estão desnaturadas e depende exclusivamente da sequência de nucleotídios
- A desnaturação e a renaturação do DNA, assim como a carga elétrica e a absorção UV, são propriedades dos ácidos nucleicos que são exploradas no desenvolvimento das tecnologias moleculares
- A migração de uma molécula de DNA dupla fita na eletroforese é uma relação de seu tamanho, sendo que moléculas menores migram mais rapidamente do que moléculas maiores
- A determinação de um fragmento de DNA que contenha uma sequência específica pode ser feita pelo método denominado Southern *blotting*
- O que se chama sonda de DNA é um fragmento de DNA com a sequência específica que se deseja identificar. Essa sonda, desnaturada, será ligada por complementariedade, especificamente, ao fragmento relacionado estruturalmente
- A clonagem gênica é uma das técnicas básicas da chamada tecnologia do DNA recombinante, que é especialmente importante ao se tratar de Biotecnologia
- Os plasmídios são estruturas de DNA dupla fita, presentes naturalmente em algumas bactérias e leveduras, e que transportam genes diferentes daqueles encontrados no genoma dos organismos hospedeiros
- Todos os plasmídios apresentam uma *origem de replicação* e, portanto, podem se replicar dentro das bactérias, independentemente da replicação do genoma bateriano
- As enzimas de restrição cortam o DNA dupla fita em locais específicos chamados de sítios de restrição, clivando o esqueleto de açúcar-fosfato da cadeia de DNA, formando extremidades livres. Essas extremidades dos fragmentos de DNA formados podem ser de 2 tipos: extremidades coesivas e extremidades cegas
- Os vetores, portanto, são plasmídios modificados em sua estrutura para funcionar, de maneira geral, na produção de múltiplas cópias de fragmento de DNA ou produzir um produto proteico, pela inserção de um gene de interesse
- A PCR é utilizada para se obter a amplificação seletiva de determinada região de uma molécula de DNA
- A PCR está baseada em 3 etapas de temperaturas diferentes que se repetem várias vezes, constituindo ciclos para a ação da enzima DNA polimerase
- A síntese de iniciadores é crítica para o sucesso da PCR. Tipicamente, esses iniciadores são oligonucleotídios contendo entre 18 e 24 nucleotídios, com temperatura de anelamento ou Tm semelhantes
- Em condições ideais, a partir do terceiro ciclo, quando as moléculas longas não são mais amplificadas e o produto tem o tamanho desejado, este será amplificado em escala exponencial. Assim, depois de 30 ciclos completos, haverá mais de 130 milhões de moléculas de DNA, com a sequência e o tamanho desejados, com base em uma molécula-molde
- A PCR quantitativa ou qPCR foi desenvolvida baseando-se na detecção dos produtos da PCR, de modo contínuo, durante o desenvolvimento da reação de amplificação. Por isso essa metodologia também é chamada de PCR em tempo real
- Na PCR em tempo real, pode-se determinar, de maneira bastante acurada, a quantidade de DNA presente no início da PCR por meio de comparação do número de ciclos
- Na RT-PCR, ocorre a amplificação por PCR do cDNA (DNA complementar ao mRNA) obtido pela ação da enzima transcriptase reversa
- Para a análise quantitativa destes, os pesquisadores desenvolveram a RT-PCR quantitativa (qRT-PCR), na qual a PCR em tempo real está acoplada à reação da transcriptase reversa. Assim, pela determinação da quantidade de transcritos produzidos, pode-se observar alterações na expressão gênica e compará-las
- Os microarranjos são arranjos de moléculas de DNA predeterminados, em geral oligonucleotídios de 20 a 25 mer, ligados quimicamente e colocados em linha em uma superfície que serve de base
- O conhecimento da sequência correta de nucleotídios de um fragmento de DNA, de um gene ou mesmo do genoma de um organismo se tornou etapa indispensável na pesquisa em genética
- O método de Maxam-Gilbert consiste na modificação química e diferencial das purinas – A e G e pirimidinas – C e T de moléculas de DNA marcadas radioativamente, seguida de

clivagem, resultando na formação de fragmentos de DNA de diferentes tamanhos

- O método de Sanger, método enzimático ou método de terminação de cadeia, se baseia na síntese de uma fita de DNA com base no DNA-molde, a ser sequenciado, utilizando no meio reacional os didesoxinucleotídios, que são nucleotídios modificados que promovem a terminação prematura da nova cadeia
- Os estudos em larga escala se baseiam no sequenciamento de milhares de segmentos de DNA ao mesmo tempo, o que promove a possibilidade de produzir uma enorme quantidade de sequências em menor tempo.

AUTOAVALIAÇÃO

10.1 Comente a seguinte sentença: "O desenvolvimento de técnicas moleculares de estudo do DNA pode auxiliar no diagnóstico das doenças genéticas humanas.".

10.2 Descreva as 3 principais propriedades físico-químicas dos ácidos nucleicos exploradas nas diversas técnicas utilizadas em biotecnologia.

10.3 Defina temperatura de anelamento de um oligonucleotídio em termos de temperatura de fusão ou Tm.

10.4 Descreva sucintamente eletroforese de ácidos nucleicos.

10.5 Quais são as principais etapas envolvidas na tecnologia de clonagem gênica?

10.6 O que são endonucleases de restrição? Dê exemplos.

10.7 Descreva as principais etapas da reação em cadeia da polimerase (PCR).

10.8 Como você diferencia a PCR convencional da PCR em tempo real?

10.9 O que é cDNA? Como essa molécula é produzida?

10.10 Em que se baseia a técnica de Sanger para o sequenciamento de DNA?

10.11 Na sua opinião, como o chamado sequenciamento de última geração influenciará os estudos em genética?

BIBLIOGRAFIA

BROWN, T.A. *Gene cloning and DNA analysis.* 6th ed., England: John Wiley & Sons, 2010.

CAMPBELL, N.A.; REECE, J.B. *Biologia.* 8ª ed., Porto Alegre: Artmed, 2010.

HAWKINS, R.D.; HON, G.C.; REN, B. Next-generation genomics: an integrative approach. *Nature Genetics Reviews* 2010; 11: 476-486.

KWOK, S.; HIGUCHI, R. Avoiding false positives with PCR (review). *Nature* 1989; 339: 237-238.

MAXAM, A.M.; GILBERT, W. A new method of sequencing DNA. *Proc. Natl. Acad. Sci.* 1977; 74: 560-564.

MESELSON, M.; YUAN, R. DNA restriction enzyme from *E. coli.* Nature 1968; 217: 1110-1114.

MULLIS, K.; FALOONA, F. Specific synthesis of DNA *in vitro* via polymerase catalyzed chain reaction. *Methods Enzimol.* 1987; 155: 335-350.

MULLIS, K.B.; FERRÉ, F.; GIBBS, R.A. PCR, the polymerase chain reaction. EUA: Birkhauser Boston, 1994.

SAMBROOK *et al.* Molecular cloning: a laboratory Manual. 3th ed., EUA: Cold Spring Harbor Laboratory, 2001.

SANGER, F.; COULSON, A.R. A rapid method for determining sequences in DNA by primed with DNA polymerase. *J. Mol. Biol.* 1975; 94: 444-448.

SNUSTAD, D.P.; SIMMONS, M.J. *Fundamentos de genética.* 4ª ed., Rio de Janeiro: Guanabara Koogan, 2008.

SOUTHERN, E.M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 1975; 98: 503-517.

STRACHAN, T.; READ, A.P. *Genética molecular humana.* 2ª ed., Porto Alegre: Artmed, 2002.

SUTCLIFFE, J.G. pBR322 and the advent of rapid DNA sequencing. *Trends Biochem. Sci.* 1995; 20:87-90.

WATSON, J.D. *et al. DNA recombinante, genes e genomas.* 3ª ed., Porto Alegre: Artmed, 2009.

Glossário

A definição dos termos que constam deste glossário nem sempre consigna sua acepção mais estrita, como ocorreria em um dicionário. Muitas vezes o significado das palavras está em acordo com o contexto do trecho do livro em que o termo aparece; em razão disto, as diversas acepções estão separadas por uma barra em pé (|).

A

Absorbância – Quantidade de intensidade de luz que uma amostra absorve. A absorbância de uma solução é função linear da concentração de um composto

Ácido nucleico – Macromolécula formada por subunidades denominadas nucleotídios que armazena a informação genética. Existem dois tipos de ácidos nucleicos: DNA e RNA

Adaptabilidade – Componente hereditário do fenótipo que confere ao organismo vantagem em termos de sobrevivência e de reprodução em ambientes específicos

Adutos de DNA – Segmento da molécula do DNA ligado covalentemente a um composto químico

Alelo – Forma alternativa de um gene que ocupa um *locus* cromossômico específico | Uma das sequências de DNA possíveis que um gene pode apresentar

Alergia – Resposta imunológica exacerbada dirigida a substâncias estranhas ao organismo

Aminoácido essencial – Tipo de aminoácido que tem de estar presente na dieta humana, uma vez que não existem vias para sua biossíntese. Caso esse tipo de aminoácido não seja fornecido, todas as proteínas que apresentam o aminoácido em sua estrutura não serão sintetizadas

Aminoacilação – Processo prévio ao início da tradução que liga um aminoácido a seu tRNA correspondente

Amplificação gênica – Tipo de mutação que leva à ocorrência de múltiplas cópias de um gene em um genoma

Anelamento – Ligação de oligonucleotídios iniciadores ou *primers* ao DNA-alvo, por complementariedade

Aneuploidia – Alteração cromossômica que resulta em um número de cromossomos que não é múltiplo exato do conjunto haploide. As formas mais comuns de aneuploidia são as trissomias (adição de um cromossomo extra) e as monossomias (perda de um cromossomo de um par de homólogos) | Alterações no número normal de cromossomos em uma célula

Aneussomia segmentar – Perda de um segmento cromossômico em um dos membros de um par de homólogos, levando à haploinsuficiência de genes contidos no segmento cromossômico perdido

Antecipação – Manifestação mais precoce e aumento da gravidade de algumas doenças genéticas a cada geração

Antecipação genética – Aumento da severidade da doença e início progressivamente mais precoce ao longo de sucessivas gerações

Anticódon – Sequência de 3 bases em uma molécula de tRNA que pareia com um códon específico no mRNA

Antígenos – Estruturas capazes de induzir a resposta imunológica específica

Aparato de transcrição basal – Complexo de fatores de transcrição, RNA polimerase e demais proteínas montados no promotor que são capazes de iniciar a transcrição em níveis basais (mínimos)

Atenuação – Tipo de controle da expressão gênica em alguns óperons bacterianos, nos quais a transcrição é finalizada prematuramente antes da RNA polimerase atingir os genes estruturais

ATP – Da classe dos nucleotídios, a adenosina 5′-trifosfato (ATP) é uma molécula rica em energia, em virtude das ligações fosfato que, quando rompidas ou criadas, são utilizadas na transferência de energia química nas reações do metabolismo

Autofecundação – Fusão de gametas de um mesmo indivíduo

Autossomo – Qualquer cromossomo nuclear diferente dos cromossomos sexuais X e Y. Existem 22 pares de autossomos nas células somáticas humanas

Azoospermia – Ausência total de espermatozoides no sêmen

B

Bandeamento cromossômico – Grupo de técnicas citogenéticas promovem a coloração diferenciada de cromossomos ou regiões cromossômicas, possibilitando a identificação de alterações cromossômicas

Base nitrogenada – Molécula heterocíclica formada por átomos de carbono e de nitrogênio. Existem dois tipos de bases nitrogenadas: pirimidinas (T, C e U) e purinas (A e G)

Bioinformática – Conjunto de metodologias associadas ao uso de computadores, programas e modelos matemáticos em processamento, integração e ordenação de dados biológicos

Biologia molecular – Ramo da biologia que estuda a estrutura e a função do material genético, utilizando técnicas em química, bioquímica e

microbiologia para a investigação de processos como a replicação celular e a expressão gênica

Bolha de replicação – Estrutura semelhante a uma bolha que se forma na região onde o DNA está sendo duplicado, em função da separação das duas fitas da molécula

Bolha de transcrição – Segmento da molécula de DNA que foi desenrolado para expor uma sequência unifilamentar, a qual será transcrita em RNA

C

Cadeia polipeptídica – Longas sequências de aminoácidos unidos por ligações peptídicas

Camundongos knockout – Linhagem de camundongo gerada pela anulação de determinado gene ou grupo de genes por técnicas de biologia celular e molecular

Câncer – Conjunto de doenças caracterizadas por crescimento celular descontrolado, morfologia do tecido alterada e capacidade de migração das células para outros órgãos

Capacidade de revisão – Habilidade das polimerases de remover e substituir nucleotídios incorporados incorretamente

Capeamento 5′ – Modificação na extremidade 5′ de um mRNA eucariótico pela adição de um nucleotídio modificado quimicamente, importante para o trajeto do núcleo ao citoplasma e para a ligação dos ribossomos durante a tradução

Característica multifatorial – Aquela que resulta da combinação de vários fatores genéticos e ambientais

Característica poligênica categórica – Fenótipo poligênico que apresenta um número definido de classes fenotípicas e não uma infinidade, como observado nas características quantitativas contínuas

Característica poligênica contínua – Característica quantitativa que exibe variação fenotípica contínua, ou seja, os fenótipos se distribuem gradualmente por um espectro sucessivo de medidas

Carboidratos – Classe de moléculas também conhecidas como glicídios ou açúcares, que podem ser classificados segundo quantidades de unidades de glicídios em poli-, tri-, di- ou monossacarídios

Cariótipo – Constituição cromossômica de um indivíduo, na qual os cromossomos são ordenados aos pares e em ordem decrescente de tamanho

Casamento consanguíneo – União entre pessoas aparentadas, ou seja, que têm um ou mais ancestrais em comum

cDNA – DNA complementar formado a partir de uma sequência de RNA

Célula somática – Refere-se a qualquer célula de um organismo diferente das germinativas ou dos gametas. As células que formam os tecidos do corpo são somáticas

Células auxiliares de memória – Subtipos de linfócitos B ou T produzidos após a resposta imunológica primária

Células citotóxicas naturais – Células conhecidas como *natural killers* ou NK, são linfócitos citotóxicos especializados em destruir células infectadas por vírus ou cancerosas

Células eucarióticas somáticas – Células diploides que compõem o organismo

Células fagocitárias – Células que têm capacidade de fagocitar, ou seja, podem emitir prolongamentos celulares que vão englobar partículas sólidas e bactérias, promovendo sua destruição

Células-tronco – Células com grande potencial de replicação, não diferenciadas, que irão produzir outras células semelhantes ou células comprometidas na formação de tecidos

Centimorgan – Parâmetro utilizado para medir a distância genética entre dois *loci*. Dois *loci* estarão a 1 cM um do outro se a frequência de recombinação entre eles for de 1%

Centrômero – Constrição primária de um cromossomo, que o separa em dois braços. É a posição na qual o cromossomo se liga ao fuso durante a divisão celular, por meio dos cinetócoros | Constrição primária existente nos cromossomos eucarióticos que mantém as cromátides irmãs unidas e delimita os braços cromossômicos (braço curto e braço longo)

Centrossomo – Região na qual um par de centríolos organiza a formação dos microtúbulos do fuso durante a divisão celular

Chaperonas – Proteínas que auxiliam na dobradura de outras proteínas

Chip de DNA – Arranjo ordenado de sequências de DNA fixadas sobre um microssuporte sólido usado para análises de ligação com uma sequência-teste (DNA ou RNA)

Choque anafilático – Resposta imunológica exagerada e aguda no indivíduo quando este entra em contato pela segunda vez com a substância estranha que promove reações patológicas importantes como inconsciência, convulsões e edema de glote

Ciclo celular – Evento celular cíclico que envolve a divisão celular e a interfase (período entre duas divisões)

Cinetócoro – Estrutura presente no centrômero, formada pela reunião de proteínas, que serve como ponto de ligação dos cromossomos às fibras do fuso

Citocinas – Moléculas existentes no meio extracelular que atuam na transmissão de sinais entre as células

Citocinese – Divisão do citoplasma que resulta nas células-filhas durante a mitose ou a meiose | Estrangulamento da membrana celular da célula-mãe na telófase, por meio da formação de um anel contrátil, que torna possível a separação final das células-filhas

Citogenética – Estudo dos cromossomos

Cloroplasto – Organela citoplasmática de plantas e algas, que contém DNA e clorofila. No cloroplasto, ocorre a fotossíntese; também se autoduplica

Código genético – Conjunto de regras que governa a tradução de uma molécula de mRNA em uma proteína

Codominância – Característica em que os efeitos de dois alelos distintos presentes em um *locus* no par de cromossomos homólogos são observados no fenótipo

Codominante – Quando os efeitos de 2 alelos distintos presentes em um *locus* no par de cromossomos homólogos são observados no fenótipo

Códon – Sequência de 3 nucleotídios adjacentes capazes de especificar um aminoácido durante a tradução

Códon de término – Códons que sinalizam para a finalização da tradução de um polipeptídio, sendo representados pelas sequências de base UAA, UAG e UGA

Coeficiente de endocruzamento – Probabilidade de uma pessoa ter dois alelos iguais por herança

Coeficiente de seleção (s) – Medida que expressa a intensidade relativa da seleção natural contra diferentes fenótipos em virtude das diferenças adaptativas que eles conferem aos organismos da população

Complexo de silenciamento induzido por RNA – Conjugação de proteínas à uma molécula de miRNA ou siRNA que possibilita a regulação pós-transcricional negativa de moléculas de mRNA-alvo

Complexo principal de histocompatibilidade ou MHC – Família de genes que codificam diferentes proteínas de superfície celular, fornecendo identidade às células de um mesmo organismo e fazendo parte do processo de apresentação de antígenos

Complexo sinaptonêmico – Arranjo complexo altamente conservado, composto por RNA e proteínas, que se forma entre os cromossomos homólogos, durante a prófase da primeira divisão meiótica

Congênito – Presente ao nascimento

Congressão – Deslocamento dos cromossomos em direção à porção central entre os dois polos da célula durante a divisão celular

Conjugação – Processo pelo qual algumas bactérias podem trocar material genético, principalmente plasmídios

Consanguinidade – Relação de parentesco

Constrição primária – Região de intensa espiralização presente em todos os cromossomos viáveis. Representa o mesmo que centrômero

Constrição secundária – Segunda região de intensa espiralização, observada nos cromossomos acrocêntricos. Nessas regiões, situam-se os genes que produzem RNA ribossômico

Conversão gênica – Permuta genética não recíproca natural, na qual uma sequência de DNA é modificada para se tornar idêntica à sequência que foi utilizada como molde (doadora). É um processo observado no cromossomo Y para preservar a função de seus genes ao longo da evolução

Corpúsculo de Baar – Estrutura condensada, altamente corável, que é encontrada na maioria das células de mamíferos e representa o cromossomo X inativo

Cromátide – Filamento de cromatina que se forma pela autoduplicação do DNA. Duas cromátides unidas pelo centrômero formam o cromossomo

Cromatina – Complexo formado pela interação entre DNA, RNA, proteínas histonas e não histonas

Cromossomo – Estrutura que contém a informação genética. Os cromossomos de eucariotos são estruturas lineares formadas por DNA e proteínas associadas. O cromossomo bacteriano é constituído por uma molécula de DNA circular

Cromossomo acêntrico – Cromossomo desprovido de centrômero, que é instável durante as divisões celulares, por não poderem se ligar ao fuso

Cromossomo acrocêntrico – Tipo de cromossomo no qual o centrômero está localizado próximo a uma das extremidades. Os braços curtos desses cromossomos (13, 14, 15, 21 e 22) têm uma segunda região de constrição e transportam genes para RNA ribossômico

Cromossomo dicêntrico – Cromossomo com dois centrômeros

Cromossomo em anel – Cromossomo estruturalmente anormal, no qual os telômeros de ambos os braços foram deletados e as duas extremidades se uniram, originando uma conformação circular, ou seja, em anel

Cromossomo Filadélfia – Cromossomo 22 anormalmente curto em virtude da translocação cromossômica que envolve os cromossomos 9 e 22 presentes na leucemia mieloide crônica

Cromossomo metacêntrico – Tipo de cromossomo no qual o centrômero está localizado no centro do cromossomo, gerando braços de tamanhos semelhantes

Cromossomo submetacêntrico – Tipo de cromossomo no qual o centrômero está um pouco deslocado para uma das extremidades, gerando um braço curto e um longo

Cromossomo telocêntrico – Tipo de cromossomo no qual o centrômero está localizado exatamente em uma das extremidades, de maneira que o cromossomo só apresenta um braço

Cromossomos homólogos – Par de cromossomos com diferentes origens parentais (um é materno e o outro paterno) que se alinham durante a meiose I e são capazes de trocar material genético entre eles. Os cromossomos sexuais X e Y, ao contrário dos autossomos, apresentam homologia apenas em suas extremidades

Crossing over – Permuta recíproca de segmentos gênicos entre cromátides não irmãs de cromossomos homólogos

Cruzamento aleatório (ou panmixia) – Reprodução aleatória entre os organismos de uma população. Corresponde à associação ao acaso de genótipos

■ D

Deleção gênica – Tipo de mutação no qual um segmento de DNA contendo um ou mais genes são perdidos

Deleções cromossômicas – Perdas de um segmento cromossômico em decorrência de quebras transversais

Deme – População local de organismos que entrecruzam

Deriva genética – Variação aleatória nas frequências gênicas em populações pequenas, causadas por erros de amostragem

Desnaturação – Processo pelo qual moléculas biológicas perdem sua estrutura quando submetidas a condições diferentes das que estavam no modo funcional. As condições alteradas podem ser temperatura, pH ou pressão

Desoxirribose – Molécula de açúcar com cinco átomos de carbono que forma o DNA

Desvio padrão (s) – Medida estatística, representada pela raiz quadrada da variância, que expressa a variação de um conjunto de valores ao redor da média

Dextrógira – Refere-se ao giro à direita da dupla hélice de DNA

Dictióteno – Período em que um ovócito primário permanece desde o período do nascimento até a sua completa diferenciação (ovulação)

Diploide – O que tem duas cópias de cada cromossomo. As células somáticas humanas são diploides, ou seja, 2n

Disjunção – Separação das cromátides-irmãs para polos opostos da célula durante a divisão celular

Dispermia – Fertilização de um ovócito por dois espermatozoides

Dissomia uniparental – Erro que leva uma célula ou organismo com número cromossômico normal a ter uma mesma origem parental incorreta para um par de cromossomos em particular | Refere-se ao mecanismo que faz um indivíduo herdar ambos os cromossomos de um par de homólogos de um mesmo genitor e nenhuma cópia desse cromossomo do outro genitor

Divisão equacional – Segunda divisão meiótica que mantém o número de cromossomos da célula-mãe para as células-filhas. Cada cromossomo das células-filhas exibirá uma das cromátides presentes na célula parental

Divisão reducional – Primeira divisão meiótica que reduz à metade o número de cromossomos da célula-mãe diploide para as células-filhas haploides

DNA helicase – Enzima que desenrola e separa a dupla hélice de DNA na forquilha de replicação

DNA ligase – Enzima que catalisa a união de polinucleotídios. Ela estabelece a ligação covalente entre a extremidade 5′ de um fragmento de DNA e a extremidade 3′ de outro fragmento. Responsável pela união dos fragmentos de Okazaki

DNA polimerase – Enzima com atividade de produzir novas moléculas de DNA com base em uma molécula de DNA preexistente utilizada como molde | Enzima responsável pela replicação do DNA. Todas as DNA polimerases replicativas adicionam desoxirribonucleotídios à fita em formação no sentido 5′→ 3′

DNA repetitivo – Sequência de DNA presente em múltiplas cópias no genoma

DNA satélite – Fração do DNA genômico formada por sequências altamente repetidas que pode ser separada do DNA total por centrifugação

DNA topoisomerase – Enzima que promove quebras transitórias no DNA, auxiliando seu desenrolamento e impedindo a super-helicoidização na região anterior à forquilha de replicação

Doenças autoimunes – Grupo de doenças provocadas por disfunção do sistema imunológico, que passa a produzir resposta imune contra elementos próprios do organismo

Dominância – Característica expressa fenotipicamente em heterozigotos para o alelo determinante do traço

Dominante – Qualquer característica que se expressa fenotipicamente em heterozigotos para o alelo determinante do traço

Dupla hélice – Estrutura química do DNA (duas fitas pareadas e enroladas formando uma hélice)

Duplicação cromossômica – Alteração estrutural que envolve a adição de uma cópia de um determinado segmento cromossômico

Duplicação semiconservativa – Característica da replicação do DNA. A duplicação desta molécula origina duas novas moléculas, cada uma formada por uma fita de DNA molde e uma fita de DNA recém-sintetizada

Duplo heterozigoto – Indivíduo que apresenta heterozigose em cada 1 de 2 *loci* distintos

■ E

Edição do RNA – Processo natural raro representado pela modificação da sequência de um transcrito de RNA, seja pela adição de grupamentos químicos, seja pela alteração de nucleotídios originalmente presentes

Efeito do fundador – Diferença na frequência de determinado genótipo originado de um pequeno número de indivíduos portadores, ditos fundadores, em uma população isolada, diferindo da população de origem | Refere-se ao estabelecimento de uma população, com base em um pequeno grupo de organismos, cujos genótipos contêm apenas uma fração dos alelos da população original. Constitui causa de deriva genética

Elemento de transposição (elemento genético transponível) – Sequência de DNA, presente em procariotos e eucariotos, capaz de se mover de um local para outro no genoma

Eletroforese – Método bioquímico que tem como principal objetivo a separação analítica ou preparativa de biomoléculas com base nas propriedades elétricas e de massa destas

Encadeamento (splicing) – Parte do processo de maturação do pré-mRNA. Refere-se à retirada dos íntrons e união dos éxons, originando uma molécula de mRNA constituída somente por regiões codificantes

Endocitose – Processo pelo qual a célula pode internalizar substâncias ou organismos externos pela captação destas e envolvimento com a membrana celular

Endocruzamento (endogamia) – Tipo de reprodução preferencial que ocorre entre organismos geneticamente aparentados

Endonuclease – Grupo heterogêneo de enzimas que clivam ligações entre nucleotídios nas fitas de DNA ou de RNA

Endossomos – Organelas formadas pela ação da endocitose

Enzima de restrição – Enzima de origem bacteriana, também denominada endonuclease de restrição, que age clivando o DNA dupla fita em sequências de bases específicas que são os sítios de restrição

Enzimas – Proteínas ou RNA com características catalíticas, ou seja, que promovem reações químicas em sistemas vivos

Enzimas modificadoras de DNA – Enzimas com atividades ligadas à manipulação de DNA, de maneira a modificar a molécula de DNA *in vitro*, ao cortar, ligar, encurtar, alongar, copiar em DNA ou RNA, adicionar e remover grupos químicos

Epigenética – Estudo das modificações hereditárias na regulação de um gene que não são atribuíveis a mudanças em sua sequência de DNA

Epistasia – Refere-se a interações não recíprocas entre produtos de genes não alélicos, de modo que um gene em um *locus* pode interferir na expressão de outro gene situado em outro *locus*

Equilíbrio de Hardy-Weinberg – Condição de um estado de equilíbrio genético, na manutenção das frequências alélicas e genotípicas de uma geração para outra em grandes populações com reprodução aleatória, que não sofrem a ação de mutação, migração ou seleção natural. Também chamada lei, princípio ou modelo de Hardy-Weinberg

Erro de amostragem – Desvio da proporção esperada, decorrente do acaso, em virtude do número de eventos ser pequeno

Erros inatos do metabolismo – Termo utilizado para descrever o conjunto de doenças genéticas do metabolismo

Espectrofotometria – Método óptico de medida de absorção ou transmissão de luz do espectro visível ou invisível. Esse método é muito utilizado para identificação e quantificação de compostos

Espermatozoide – Célula germinativa masculina haploide ou gameta masculino

Estrutura quaternária – Nível de organização de uma proteína cuja configuração é constituída pelo arranjo tridimensional de duas ou mais cadeias polipeptídicas

Eucarioto – Organismo cujas células têm núcleo verdadeiro, ou seja, delimitado por uma membrana nuclear

Eucromatina – Parte da cromatina mais distendida, levemente corada durante a interfase | Região da cromatina que apresenta estrutura mais relaxada e, por conseguinte, maior concentração de genes em atividade. A configuração descondensada faz que essa região absorva pouca quantidade de corante

Euploidia – Alteração cromossômica que resulta em um número de cromossomos que é múltiplo exato do conjunto haploide

Éxon – Região codificante de um gene eucariótico, que, após transcrição e tradução, estará representada na cadeia polipeptídica codificada pelo gene | Segmento de DNA que está representado no mRNA maduro e que participará como molde na síntese proteica

Expressividade variável – Variabilidade observada nos efeitos fenotípicos produzidos por um mesmo gene em diferentes indivíduos

■ F

Família gênica – Agrupamento de genes relacionados funcionalmente e que ocupam diversos *loci* no genoma. Acredita-se que os genes de uma família tenham surgido de um único gene ancestral, principalmente pela similaridade de suas sequências nucleotídicas

Fase aberta de leitura ou ORF – Sequência de DNA de um gene que é traduzida em uma proteína

Fator de transcrição – Proteínas que se ligam a sítios específicos no genoma e que ativam a transcrição

Fenocópia – Traço presente em um indivíduo com genótipo normal, causado por um fator ambiental que se assemelha a uma condição genética determinada por um genótipo específico

Fenótipo – Característica observada em um organismo, determinada pela interação de seu genótipo com o ambiente em que se expressa

FISH – Hibridação *in situ* com fluorescência

Fita atrasada ou tardia (lagging strand) – Fita de DNA sintetizada de modo descontínuo, na direção oposta à abertura da forquilha, por meio de curtos fragmentos denominados fragmentos de Okazaki

Fita líder (leading strand) – Fita de DNA nascente polimerizada de modo contínuo, no mesmo sentido da abertura da forquilha de replicação

Fluxo gênico – Transferência de alelos de uma população para outra em virtude da migração de organismos férteis

Forquilha de replicação – Bifurcação formada na extremidade da bolha de replicação, em função do desenrolamento da dupla hélice, na qual o DNA está sendo duplicado

Fragmento de Okazaki – Curta sequência de nucleotídios sintetizada durante a replicação descontínua da fita tardia de DNA

Frequência alélica (ou gênica) – Refere-se à proporção dos alelos de um gene em uma população

Frequência genotípica – Refere-se à proporção de cada genótipo em um *locus* específico entre os membros de uma população

■ G

Gameta – Célula da linhagem reprodutiva (espermatozoide ou ovócito). Apresenta número haploide de cromossomos (em humanos, 23 cromossomos)

Gargalo genético – Redução drástica no tamanho da população, com consequente perda da variabilidade genética, que leva à deriva genética

Gêmeos dizigóticos – São aqueles originados de uma ovulação materna dupla e de dois eventos independentes de fertilização, sendo geneticamente tão semelhantes quanto dois irmãos não gêmeos

Gêmeos monozigóticos – São aqueles originados da divisão e separação de um único zigoto e, por essa razão, possuem genótipos idênticos

Gene – Unidade física e funcional da hereditariedade, formada por um segmento de DNA específico (ou de RNA, em alguns vírus)

Genes deletérios – Genes mutados que não desempenham adequadamente suas funções na célula, causando distúrbios ao organismo

Genes housekeeping – Genes que regem funções celulares básicas para o organismo e são expressos em níveis basais constantes

Genes supressores tumorais – São genes cujos produtos proteicos participam do controle negativo do ciclo celular

Genética molecular – Ramo da genética que estuda as bases moleculares da hereditariedade e das patologias humanas de origem genética

Genoma – Informação genética total presente na célula de um organismo

Genômica – Ramo da genética que descreve e analisa a constituição do genoma de uma espécie. Seu desenvolvimento é muito recente e depende de genomas completamente sequenciados e de ferramentas de bioinformática

Genótipo – Constituição genética de uma célula ou de um organismo. Mais especificamente, o termo é usado em referência aos alelos presentes em um *locus* cromossômico particular

Genótipos recombinantes – Novas combinações genotípicas distintas dos genótipos parentais

Grau de liberdade (gl) – Índice associado aos dados independentes que são utilizados na estimativa de um parâmetro. Em relação ao qui-quadrado para testar o equilíbrio de Hardy-Weinberg, *gl* corresponde ao número de classes dos dados menos o número de parâmetros estimados menos um

■ H

Haploide (n) – Célula que tem um membro de cada par de cromossomos

Haploinsuficiência – Quando a presença de apenas uma cópia normal de um alelo não é suficiente para prevenir a manifestação de um fenótipo anormal em virtude da perda de função do outro alelo

Haplótipo – Conjunto de genes ou variantes alélicas localizadas em um mesmo cromossomo que são transmitidos juntos à próxima geração. Os haplótipos são vastamente utilizados para seguir genealogias e populações, além de serem importantes no mapeamento cromossômico

Heme – Grupamento químico constituído por um átomo de ferro localizado no centro de um anel orgânico heterocíclico chamado de porfirina

Hemizigoto (hemizigose) – Termo utilizado principalmente para os genes no cromossomo X que estão apenas em uma cópia nos homens em relação às mulheres | Aquele que tem somente um alelo em determinado *locus* nas células diploides. Os homens são hemizigotos para genes ligados ao X por terem apenas um cromossomo X

Hemoglobina – Proteína presente nas hemácias, cujas funções principais são fazer o transporte de oxigênio dos pulmões para os tecidos e manutenção do pH interno constante

Hemoglobinopatias – Conjunto de doenças resultantes de alterações patogênicas na proteína hemoglobina, presente nas hemácias, cuja principal função é transportar oxigênio dos pulmões para os tecidos

Herança digênica – Tipo de herança oligogênica, na qual o fenótipo resulta de combinações de alelos em 2 genes

Herança holândrica – Padrão de herança de genes ligados ao cromossomo Y

Herança materna (ou matrilinear) – Padrão de herança mitocondrial no qual a transmissão da informação genética contida no mtDNA ocorre apenas pela mulher

Herança mendeliana – O mesmo que herança monogênica. Refere-se aos padrões de transmissão gênica que seguem as leis de Mendel

Herança monogênica – O mesmo que herança mendeliana. Refere-se aos padrões de transmissão de alelos de um único gene, presente no genoma nuclear de um organismo diploide

Herança multifatorial – Refere-se à herança de fenótipos determinados por interações entre vários fatores genéticos e ambientais

Herança oligogênica – Refere-se à herança de fenótipos que resultam da ação combinada de poucos genes (em geral, 2 a 15)

Herança poligênica – Refere-se à herança de características fenotípicas cuja expressão depende da ação conjunta de vários genes, cada um produzindo um pequeno efeito sobre o traço | Tipo de herança em que o fenótipo é resultado da combinação da expressão de variantes de um conjunto de genes

Herança quantitativa – O mesmo que herança poligênica

Herdabilidade – Mede a contribuição relativa da variância genética na variabilidade fenotípica total de uma condição multifatorial em uma população

Herdabilidade em sentido amplo (H^2) – Proporção da variância fenotípica total de um traço quantitativo que pode ser atribuída à variância genotípica total e não apenas aos efeitos aditivos dos poligenes

Herdabilidade em sentido restrito (h^2) – Proporção da variância fenotípica total de um traço quantitativo decorrente da variância genética aditiva

Heredograma – Diagrama usado para mostrar a transmissão familiar de um traço hereditário

Heterocromatina – Região da cromatina que se cora fortemente até mesmo na intérfase, por apresentar um aspecto condensado. A heterocromatina pode ser constitutiva (p. ex., regiões centroméricas) ou facultativa (um dos cromossomos X na mulher)

Heterocromatina constitutiva – Cromatina que aparece sempre de forma heterocromática, presente nos centrômeros

Heterocromatina facultativa – Cromatina que pode existir como heterocromatina ou eucromatina, dependendo do estado da célula

Heterodissomia – Evento de dissomia parental, no qual os dois alelos ou cromossomos com a mesma origem parental são diferentes

Heterogeneidade alélica – Refere-se à ocorrência de um mesmo fenótipo causado por diferentes mutações em um mesmo gene

Heterogeneidade de locus – Refere-se à ocorrência de um mesmo fenótipo causado por mutações em genes diferentes

Heterogeneidade genética – Produção de fenótipos iguais por mecanismos genéticos diferentes

Heteroplasmia – Presença de, pelo menos, dois tipos de sequências de DNA mitocondrial em um indivíduo

Heterozigoto (heterozigose) – Organismo com dois alelos diferentes em determinado *locus* no par de cromossomos homólogos

Heterozigoto composto – Indivíduo que apresenta um traço autossômico recessivo em decorrência de duas mutações diferentes em um mesmo gene

Hipótese de Lyon – Hipótese formulada por Mary Lyon em 1961 que preconiza que um cromossomo X em cada célula feminina é desligado de modo aleatório para garantir uma compensação de dose entre os sexos feminino e masculino

Histona – Grupo de proteínas básicas, evolutivamente conservadas, que se ligam ao DNA formando a cromatina

Holoenzima – Forma de uma enzima multimérica, na qual todos os componentes polipeptídicos estão presentes

Homeostase – Manutenção das condições constantes do meio interno de um organismo

Homoplasmia – Presença de somente um tipo de sequência de DNA mitocondrial em um indivíduo

Homozigoto (homozigose) – Organismo com dois alelos iguais em determinado *locus* no par de cromossomos homólogos

■ I

Ilhas CpG – Segmento da molécula de DNA que é rico em dinucleotídios CG ligados covalentemente e é frequentemente encontrado nas regiões promotoras dos genes

Imprinting genômico – Marcação epigenética diferencial que ocorre na cromatina e acarreta padrões distintos de expressão gênica

Imprinting parental – Determinação da expressão de um gene pela origem parental

Imunização ativa – Tipo de resposta imunológica que diz respeito à produção de anticorpos ou linfócitos T pelo organismo quando este entra em contato com o agente infeccioso

Imunização passiva – Tipo de resposta imunológica que diz respeito à produção de anticorpos ou linfócitos T por um organismo e transferida para outro

Imunoglobulinas – Proteínas produzidas pelos linfócitos B diferenciados em plasmócitos, que atuam ao se ligar a determinantes antigênicos e que ajudam em sua eliminação; são também denominados anticorpos

Iniciação – Primeira etapa para a formação do câncer, na qual carcinógenos produzem modificações no material genético, tornando a célula uma célula alterada

Intercinese – Período de relaxamento da estrutura cromossômica entre a primeira e a segunda divisão meiótica, na qual não ocorre replicação do DNA

Interfase – Período do ciclo celular entre duas divisões celulares | Período entre duas mitoses sucessivas

Interferons – Grupo de proteínas produzidas por células humanas teciduais ou linfócitos que têm a propriedade de induzir resistência viral às células não infectadas

Íntron – Região gênica não codificante, que se intercala às regiões codificadoras e que é removida durante o *splicing* do mRNA, não estando representada no polipeptídio codificado pelo gene | Segmento de DNA que é transcrito, mas posteriormente removido do mRNA gerado, e que não servirá de molde na síntese proteica

Inversão – Alteração que envolve a excisão de um segmento cromossômico, seguida de um movimento de 180° e a reintrodução do segmento em ordem invertida. Pode ser paracêntrica ou pericêntrica

Ionização – Processo químico no qual são formadas estruturas químicas eletricamente carregadas pela perda ou pelo ganho de elétrons

Isodissomia – Evento de dissomia parental, no qual os dois alelos ou cromossomos com a mesma origem parental são iguais

■ L

LDL – Sigla que designa lipoproteína de baixa densidade. É a lipoproteína formada pela parte proteica, ou apolipoproteína, e moléculas de colesterol. As moléculas de colesterol dão o caráter de baixa densidade dessa lipoproteína

Levógira – Refere-se ao giro à esquerda da dupla hélice de DNA

Ligação genética – Transmissão conjunta de dois *loci* para um mesmo gameta durante a meiose, em virtude da grande proximidade entre eles e a ausência de recombinação

Linfócitos B – Células que se originam na medula óssea e vão se diferenciar nesse mesmo local, mas, segundo estímulos, podem formar os plasmócitos, que são as células circulantes produtoras de anticorpos

Linfócitos B e T – Células especializadas participantes da resposta imune dos organismos vertebrados

Linfócitos T – Células que se originam na medula óssea e vão se diferenciar no timo, e, segundo os estímulos, vão formar diferentes tipos de linfócitos T

Linfócitos T auxiliares – Células T que respondem à apresentação do antígeno pelas células apresentadoras de antígeno ativando linfócitos B a produzir anticorpos e linfócitos T a produzir receptores específicos, induzindo tanto a resposta imunológica humoral quanto a celular

Linhagem germinativa – Tecido reprodutivo de um organismo

Lipoproteínas – Classe de complexos moleculares formados por proteínas, chamadas de apolipoproteínas, e lipídios. Elas são identificadas segundo a densidade que depende do tipo de lipídio e têm, primariamente, a função de transporte para sua metabolização no fígado

Lisossomos – Organelas celulares envolvidas no processo de digestão de compostos intra e extracelulares

Loci de característica quantitativa (QTL, quantitative trait loci) – Genes que contribuem para um traço quantitativo

Loci sintênicos – *Loci* gênicos localizados em um mesmo cromossomo, independentemente do fato de estarem ligados ou não

Locus (plural, loci) – Posição que um gene ocupa ao longo de um cromossomo. Um único *locus* pode ter diferentes alelos

Lod score – Estimativa estatística para aferir se dois *loci* estão provavelmente próximos um do outro em um mesmo cromossomo e são herdados juntos

■ M

Mapeamento físico – Representação gráfica das distâncias (em pares de bases) entre genes e de suas posições relativas em um cromossomo

Mapeamento genético – Diagrama representativo das posições relativas dos genes ao longo de um cromossomo, com base na frequência de recombinação entre eles, a qual é medida em centimorgan (cM)

Matriz extracelular – Conjunto de elementos intercelulares que promovem a interação e a sustentação das células de um tecido

Mecanismos epigenéticos – Conjunto de processos moleculares que modificam a expressão dos genes sem alterar sua sequência de nucleotídios

Média – Média aritmética

Megabase (Mb) – Unidade de comprimento. Assim, 1 Mb equivale a 1 milhão de nucleotídios

Meiose – Divisão celular realizada durante a formação dos gametas, na qual uma célula-mãe diploide dá origem a quatro células-filhas haploides, em duas etapas subsequentes (meiose reducional e meiose equacional) | Tipo de divisão celular que resulta na formação dos gametas, células haploides (n) com metade do número cromossômico da célula-mãe (2n)

Metabolismo – Conjunto de processos físico-químicos que transformam substâncias dentro de um organismo. As reações bioquímicas que fazem parte do metabolismo estão organizadas em sequência, chamadas de vias metabólicas, nas quais o produto de uma reação é utilizado como substrato para a reação seguinte

Methyl CpG binding proteins (MBP) – Grupo de proteínas capazes de se ligar a ilhas CpG e reprimir a transcrição gênica de maneira direta ou indireta

Metilação do DNA – Processo de adição de grupamentos metila no DNA por meio de enzimas DNA metiltransferases

Método de gêmeos – Refere-se ao estudo comparativo entre pares de gêmeos monozigóticos e dizigóticos, bastante utilizado na estimativa da herdabilidade

Microssatélites – Sequências de DNA que se repetem consecutivamente e contêm número de repetições variável entre diferentes indivíduos

Migração – Movimento de organismos de uma população para outra

miRNA – Pequeno RNA de interferência produzido pela clivagem de um RNA geralmente unifilamentar, cuja estrutura se dobra em forma de um grampo bifilamentar. Os miRNA se conjugam a complexos de proteínas para regular a expressão gênica em nível pós-transcricional de mRNAs-alvo

Mitocôndria – Organela citoplasmática das células eucarióticas na qual se processa a respiração celular. Ela contém DNA e se autoduplica

Mitógeno – Substância que estimula a proliferação celular

Mitose – Divisão celular realizada pelas células somáticas, na qual uma célula-mãe diploide dá origem a duas células-filhas diploides | Processo no qual ocorre a divisão da célula eucariótica, originando uma nova célula semelhante | Tipo de divisão celular que resulta na produção de duas células filhas geneticamente iguais à célula-mãe

Modelo limiar ou de propensão – Tal modelo explica a variação fenotípica quantitativa dicotômica, com base na proposição de existência de um limiar que separa os indivíduos em duas categorias: os que manifestam o fenótipo anormal e os que não manifestam

Monossomia – Alteração cromossômica numérica que surge em consequência da presença de apenas uma cópia de um cromossomo, em vez das duas habituais

Mosaicismo – Presença de duas ou mais linhagens celulares em um mesmo tecido ou em diferentes tecidos do organismo originados de um único zigoto

Mutação – Alteração na sequência de nucleotídios de um gene | Alteração ocorrida no DNA que constitui a principal fonte de variação genética

Mutação de ponto – Tipo de mutação que ocorre na molécula de DNA em que um nucleotídio é trocado por outro nucleotídio diferente

Mutação de sentido trocado – Mutação que altera o aminoácido original codificado anteriormente pelo códon

Mutação dinâmica – Expansão instável de microssatélites com o passar das gerações

Mutação direta – Mutação que transforma um alelo comum (selvagem) em um alelo raro (mutante), geralmente, deletério

Mutação esporádica – Mutação que surge nas células somáticas, sendo adquiridas durante o período de vida do indivíduo

Mutação frameshift – Alteração na sequência de nucleotídios de um gene que leva à mudança na matriz de leitura dos seus códons

Mutação gênica – Alteração na sequência de nucleotídios do DNA que constitui o gene

Mutação germinativa – Mutação presente na linhagem germinativa e que é transmitida aos descendentes do indivíduo afetado

Mutação hereditária – Mutação adquirida pela herança genética

Mutação neutra – Mutação que altera a sequência de aminoácidos de uma proteína sem ter efeito sobre sua funcionalidade

Mutação nula – Alteração na sequência de um gene que abole completamente a sua função na célula

Mutação reversa – O oposto da mutação direta. Um alelo mutante pode sofrer nova mutação e restabelecer a sequência gênica original, ou seja, o alelo selvagem | Segunda mutação que ocorre em uma mesma posição do gene em que houve uma mutação prévia, restaurando a sequência gênica original

Mutação sem sentido – Mutação que transforma uma sequência codificante em um códon de término da tradução

Mutação sense – Mutação que transforma um códon finalizador em um códon que especifica um aminoácido

Mutação silenciosa – Aquela que não altera o aminoácido original codificado por um dado códon

Mutação supressora – Segunda mutação em um local diferente no genoma que compensa os efeitos da primeira mutação

Mutações germinativas – Alterações na sequência do DNA em células que originarão os gametas

Mutações somáticas – Alterações na sequência do DNA em células do corpo que não incluem os gametas

■ N

Não disjunção – Separação inadequada dos cromossomos durante a meiose ou mitose

Norma de reação – Espectro de fenótipos que podem resultar de um genótipo específico em ambientes diversos

Northern blotting – Técnica de identificação de mRNA em uma mistura dessas moléculas por meio da separação do RNA total em eletroforese

em gel e transferência para membrana, seguida de hibridização com sondas de DNA marcadas, em fitas simples e específicas

Nucleoplasma – Parte do núcleo que exclui o nucléolo

Nucleosídio – Subunidade que forma os ácidos nucleicos constituída pela ligação de uma base nitrogenada a uma molécula de açúcar

Nucleossomo – Unidade estrutural básica da cromatina composta pelo enovelamento do DNA a um cerne de oito proteínas histona (duas de cada H2A, H2B, H3 e H4) | Unidade estrutural básica da cromatina eucariótica que consiste em um segmento de DNA enrolado ao redor de um octâmero de histonas e representa o primeiro nível de compactação da cromatina

Nucleotídio – Subunidade que forma os ácidos nucleicos, constituída por um grupo fosfato, uma pentose e uma base nitrogenada

O

Oligonucleotídios – Fragmentos curtos de DNA ou RNA, com até algumas dezenas de nucleotídios e que são, em geral, produzidos por síntese química, sob a forma de fita simples

OMIM – Catálogo *on-line* de traços humanos com herança monogênica

Oncogene – Classe de genes caracterizados pela presença de mutação ativadora ou expressão aumentada, associada ao processo de formação do câncer

Óperon – Conjunto de genes estruturais bacterianos, seu promotor comum e outras sequências reguladoras flanqueadoras que controlam a expressão dos genes estruturais

Óperon indutível – Óperon no qual a transcrição dos genes estruturais está normalmente desligada e uma mudança deve acontecer para que a transcrição seja ativada

Óperon repressível – Óperon no qual a transcrição dos genes estruturais está normalmente ligada e uma mudança deve acontecer para que a transcrição seja desativada

Origem de replicação – Sequência específica de nucleotídios presente no DNA genômico dos organismos, que marca o início da replicação do DNA | Sítios específicos na molécula de DNA, nos quais proteínas se ligam para dar início à síntese de novas moléculas de DNA

Ovócito – Célula germinativa feminina haploide ou gameta feminino

P

PCR em tempo real – Tipo de PCR que se baseia na detecção dos produtos formados pela ação da polimerase durante o processo

Penetrância – Refere-se à probabilidade que um gene apresenta de ter alguma expressão fenotípica

Penetrância completa – Refere-se aos alelos cuja expressão ao nível de fenótipo ocorre em 100% dos indivíduos portadores do gene

Pentose – Molécula de açúcar (desoxirribose no DNA e ribose no RNA) com cinco átomos de carbono presente nos ácidos nucleicos

Pequeno RNA nuclear (snRNA) – Pequenas moléculas de RNA presentes no núcleo das células eucarióticas que se ligam a proteínas e atuam no processamento do pré-mRNA

Pequeno RNA nucleolar (snoRNA) – Pequenas moléculas de RNA que participam da maturação do rRNA

Pirimidina – Base nitrogenada formada por um anel heterocíclico presente nos ácidos nucleicos (C e T no DNA, e U no RNA)

Plasmídios – Moléculas de DNA dupla-fita e circulares, extracromossômicos, presentes em alguns tipos de bactérias e que têm capacidade replicativa

Plasmídios R – Plasmídios de origem bacteriana que apresentam genes que conferem resistência a alguns antibióticos

Plasticidade fenotípica – Propriedade inerente aos organismos vivos que lhes possibilita responder de modo diferenciado a mudanças no meio ambiente

Pleiotropia – Efeitos fenotípicos múltiplos de um gene

Poliadenilação 3′ – Adição de resíduos de adenina subsequentes à extremidade 3′ de um mRNA que tem grande importância em garantir estabilidade ao mRNA

Poliadenilação alternativa – Utilização diferencial de sítios distintos de clivagem na extremidade 3′ de um transcrito de mRNA para a adição da cauda poli(A)

Poligenes – Genes envolvidos na herança quantitativa ou poligênica

Poligenes deletérios – Genes que não desempenham adequadamente suas funções na célula e causam distúrbios ao organismo

Polimorfismo – Existência de uma variante na sequência de nucleotídios de um gene que apresente frequência populacional igual ou superior a 1%

Polimorfismo de nucleotídio único (SNP, do inglês single nucleotide polymorphism) – Variação de uma única base do DNA observada no genoma de indivíduos de uma população

Polipeptídio – Cadeia de aminoácidos unidos por ligações peptídicas

Poliploide – O que tem mais de duas cópias de cada cromossomo, ou seja, triploide (3n), tetraploide (4n), ...

Polirribossomo – Molécula de mRNA com vários ribossomos a ela ligados

Ponte de hidrogênio – Ligação química fraca que se forma entre as bases nitrogenadas que pareiam entre si (adenina com timina e citosina com guanina) e mantém as duas fitas de DNA unidas

Pool gênico – Conjunto de todos os alelos compartilhados pelos organismos que constituem uma população local de entrecruzamento em determinada época

População – Grupo de organismos de uma mesma espécie, que vivem na mesma área geográfica, partilham um conjunto gênico comum e são capazes de cruzar entre si e produzir prole fértil

Porfirina – Classe de molécula orgânica formada por 4 anéis do tipo pirrol, ligados por pontes de carbono do tipo metínicas e cujo centro pode conter um íon metálico. Por causa das ligações duplas e intercaladas, essas moléculas absorvem luz visível, podendo, portanto, ser coloridas

Posição de oscilação – Pareamento entre o terceiro nucleotídio do códon do mRNA e o primeiro nucleotídio do anticódon do tRNA, no qual podem existir interações incomuns entre bases

Primase – Enzima responsável pela síntese do *primer* de RNA

Primer (oligonucleotídio iniciador) – Pequeno segmento de RNA iniciador da replicação do DNA em procariotos e eucariotos que fornece a extremidade 3′-OH, à qual a DNA polimerase adicional nucleotídios

Probando (propósito ou caso-índice) – Indivíduo que apresenta determinado traço genético e pelo qual a família começa a ser averiguada

Procarioto – Organismo unicelular, como as bactérias, que não tem núcleo delimitado por uma membrana

Progressão – Terceira etapa para a formação do câncer, cuja célula alterada já é cancerosa e tem crescimento descontrolado

Promoção – Segunda etapa para a formação do câncer, na qual agentes promotores como hormônios e fatores de crescimento vão promover e sustentar o crescimento das células iniciadas

Promotor – Sequências conservadas, às quais a RNA polimerase se liga para iniciar a transcrição de um gene

Proteína – Molécula constituída por uma ou mais cadeias de aminoácidos, com uma estrutura tridimensional específica, que lhe confere determinada função biológica

Proteína SSB – Aquela que se liga e estabiliza as fitas simples de DNA na região da forquilha de replicação para que elas possam servir de molde para a síntese das novas cadeias

Proteínas de choque térmico – Polipeptídios produzidos em muitos organismos em resposta ao calor extremo, evitando danos maiores às células

Proto-oncogene – Classe de genes associados ao controle positivo do ciclo celular que, quando apresentam mutação ou têm sua expressão aumentada, se transformam nos oncogenes

Pseudogenes – Cópias não funcionais de genes que perderam sua capacidade de gerar moléculas de RNA e/ou proteínas ao longo da evolução | Genes similares a genes funcionais, sendo que perderam sua função em virtude de mutações

Purina – Base nitrogenada formada por dois anéis heterocíclicos presente nos ácidos nucleicos (A e G no DNA e no RNA)

■ Q

Qui-quadrado (χ^2) – Teste estatístico utilizado para checar o grau de concordância dos dados observados com as previsões de uma hipótese

Quiasmas – Evidências citológicas de que o *crossing over* aconteceu durante a meiose I. Podem ser visualizados na fase de diplóteno

Quilobase (kb) – Unidade de comprimento. Assim, 1 kb equivale a 1.000 nucleotídios

Quimeirismo – Presença de duas ou mais linhagens celulares em um mesmo tecido ou em diferentes tecidos do organismo originados de pelo menos dois zigotos diferentes

Quinase – Enzima de natureza proteica que promovem a introdução de grupamentos fosfato em componentes celulares

■ R

Reação em cadeia da DNA polimerase (PCR) – Técnica para amplificação específica de DNA *in vitro*, pelo uso de iniciadores ou *primers* específicos que delimitam o segmento a ser copiado

Receptores celulares – Moléculas de natureza essencialmente proteica que interagem com substâncias externas às células, desencadeando alterações nestas

Recessivo – Qualquer característica que se manifesta fenotipicamente em homozigotos ou em heterozigotos compostos para o alelo determinante do traço

Recombinação (crossing over) – Formação de novas combinações gênicas, não encontradas nos genitores, como resultado da troca recíproca de material genético entre cromossomos homólogos durante a prófase I da meiose

Recombinação homóloga não alélica – Recombinação entre sequências homólogas desalinhadas que gera duplicações e deleções

Recomposição (splicing) do RNA – Etapa do processamento do mRNA da maioria dos genes de eucariotos, a qual consiste na excisão dos íntrons e no encadeamento dos éxons do transcrito primário com a participação do spliceossomo

Recomposição ou encadeamento alternativo (splicing alternativo) – União combinatória diferenciada de conjuntos de éxons durante a recomposição do pré-mRNA que favorece a formação de mais de um produto com base em um único gene

Região pseudoautossômica – Porção final dos braços curtos e longos dos cromossomos X e Y que têm homologia e que trocam material genético entre si durante o *crossing over* que ocorre na formação dos gametas masculinos

Região subtelomérica – Região cromossômica proximal ao telômero, com elevada concentração gênica

Regulação epigenética – Forma de regulação gênica que envolve modificações no DNA estáveis e herdáveis que não incluem mudanças na sequência de bases

Replissomo – Complexo sistema proteico-enzimático que atua de maneira coordenada em cada réplicon permitindo que a dupla hélice se desenrole e se mantenha separada durante o processo de replicação

Repressão catabólica – Controle gênico em alguns óperons bacterianos, nos quais a glicose é utilizada preferencialmente em detrimento de outros açúcares, os quais têm seu metabolismo reprimido pela presença de glicose

Retinoblastoma – Câncer que acomete a retina

Retrotransposon – Sequência de DNA móvel capaz de se transpor para diferentes locais no genoma

Ribose – Molécula de açúcar com cinco átomos de carbono que forma o RNA

Risco de recorrência – Chance de uma condição genética detectada em um ou mais membros da família aparecer em outro familiar na mesma progênie ou em gerações seguintes

RNA antissenso – RNA que tem sequência complementar a um RNA senso, tendo a capacidade de se ligar a ele, regulando sua expressão

RNA mensageiro (mRNA) – Molécula de RNA transcrita do DNA que será traduzida na sequência de aminoácidos de um polipeptídio

RNA policistrônico – Molécula de RNA transcrita de um grupo de genes contínuos capaz de promover a síntese de múltiplas proteínas

RNA ribossômico (rRNA) – Molécula de RNA que compõe as subunidades ribossômicas. O ribossomo é uma estrutura formada por duas subunidades de ribonucleoproteínas, na qual a síntese de proteínas ocorre

RNA transportador (tRNA) – Molécula de RNA que participa do processo de tradução do mRNA, transportando aminoácidos específicos durante a síntese de proteínas

■ S

Sarcoma – Tipo de câncer originado de células do tecido conjuntivo

Satélite – Região distal à constrição secundária

Seleção balanceadora – Tipo de seleção natural que favorece os heterozigotos em um *locus*, em prejuízo dos homozigotos, em razão de eles apresentarem maior adaptação

Seleção direcional – Força seletiva que favorece fenótipos em um extremo da distribuição

Seleção disruptiva – Processo seletivo que favorece fenótipos extremos, em prejuízo dos intermediários, em virtude de existir na população mais de um fenótipo bem adaptado às condições ambientais existentes

Seleção estabilizadora – Processo seletivo que favorece fenótipos intermediários, bem adaptados ao ambiente, em prejuízo dos fenótipos extremos que conferem baixa adaptação

Seleção natural – Força que opera, de maneira diferencial e não aleatória, selecionando os organismos de uma população com fenótipos mais bem adaptados a seu ambiente, ou seja, aqueles com maior êxito em relação à capacidade de sobrevivência e de reprodução

Sequência de Shine-Dalgarno – Sequência consenso encontrada na região 5′ não traduzida do mRNA de bactérias que contém o sítio de ligação do ribossomo

Sequência Kozak – Sequência de nucleotídios consenso localizada no mRNA de eucariotos, a qual tem grande importância no início do processo de tradução

Sequenciamento de DNA – Conjunto de técnicas com o objetivo de determinar a sequência de bases de nucleotídios de fragmentos de DNA

Sequências de DNA consenso – Sequências de nucleotídios comumente presentes em elementos genéticos do DNA ou RNA e que têm função específica

Sexo heterogamético – Aquele que, em relação aos cromossomos sexuais, produz 2 tipos de gametas. Por exemplo, em mamíferos, o macho é o sexo heterogamético porque produz gametas com o cromossomo X e gametas com o cromossomo Y | Sexo representado pela presença de dois cromossomos sexuais distintos, ou seja, um homem (46,XY)

Sexo homogamético – Aquele que, em relação aos cromossomos sexuais, produz apenas um tipo de gameta. Por exemplo, as fêmeas de mamíferos representam o sexo homogamético, pois todos os gametas produzidos contêm o cromossomo X | Sexo representado pela presença de dois cromossomos sexuais iguais, ou seja, uma mulher (46,XX)

Sinapse – Alinhamento entre cromossomos homólogos durante a prófase da primeira divisão meiótica

Síntese semidescontínua – Característica da replicação do DNA. Por ser contínua em uma fita e descontínua na outra, a síntese é semidescontínua

siRNA – Pequeno RNA de interferência unifilamentar produzido pela clivagem de um RNA bifilamentar que pode ter origem exógena. Atua na regulação pós-transcricional da regulação gênica ao parear com moléculas de RNA-alvo complementares

Sistema complemento – Grupo de, aproximadamente, 30 proteínas que estão no sangue e que são ativadas por uma cascata de reações bioquímicas, produzidas por interações complexas

Sistema imunológico – Conjunto de reações do organismo que mantêm o equilíbrio interno constante, constituindo a defesa deste e eliminando agentes externos ou estranhos como bactérias e vírus. Ainda, o seu mau funcionamento pode levar ao desenvolvimento das doenças autoimunes e de hipersensibilidade

snoRNA – Pequenos RNA nucleolares que participam do processamento dos rRNA eucarióticos

snRNA – Pequenos RNA nucleares que fazem parte da recomposição do pré-mRNA

Sobredominância – O heterozigoto demonstra maior adaptabilidade ao meio que ambos os homozigotos – a também chamada vantagem do heterozigoto

Solenoide – Fibra de 30 nm de diâmetro formada pelo empilhamento dos nucleossomos

Sonda de DNA – Molécula de DNA de fita simples e marcada, utilizada para a identificação específica de fragmentos-alvo de DNA ou RNA, complementares

Southern blotting – Técnica molecular utilizada na identificação de fragmentos de DNA específicos em uma mistura por meio da separação desses fragmentos em eletroforese em gel e transferência para membrana, seguida de hibridização com sondas de DNA marcadas, em fita simples e específicas

Spliceossomo – Complexo de pequenos RNA nucleares que participam da recomposição do pré-mRNA após a transcrição de genes eucarióticos

Substâncias farmacologicamente ativas – Tipo de substâncias ou moléculas orgânicas presentes nos organismos que funcionam modulando e alterando a atividade celular

■ T

Telomerase – Enzima que adiciona sequências de DNA telomérico às extremidades dos cromossomos lineares de eucariotos

Telômero – Estrutura que protege as extremidades dos cromossomos eucarióticos lineares e contribui para a integridade e a estabilidade cromossômica. É formado por curtas sequências de DNA repetidas em tandem inúmeras vezes | Região formada por sequências repetitivas em tandem (TTAGGG)n, localizadas nas extremidades dos cromossomos. Tem o papel de proteger os cromossomos de rearranjos

Temperatura de fusão ou Tm – Temperatura na qual a metade das moléculas de DNA de uma amostra está sob a forma desnaturada

Terapia gênica – Tipo de terapia associado à introdução de genes em células ou indivíduos com o objetivo de corrigir o efeito patológico de genes mutantes

tmRNA – RNA mensageiro transportador que tem propriedades de mRNA e tRNA, participando da recuperação de ribossomos que permanecem ligados à molécula de mRNA

Tradução – Síntese de um polipeptídio com base em uma molécula de RNA mensageiro

Transcrição – Síntese de uma molécula de RNA com base em um molde de DNA pela RNA polimerase

Transcrição reversa – Síntese de uma molécula de DNA com base em um molde de RNA, comum em retrovírus

Transcrito primário – Molécula de RNA eucariótica que é modificada após a transcrição até gerar uma molécula de RNA madura

Translocação cromossômica – Rearranjo de segmentos de cromossomos não homólogos

Translocação robertsoniana – Tipo especial de translocação dos cromossomos acrocêntricos, na qual ocorre fusão no centrômero ou próximo a ele, com perda subsequente dos braços curtos

Trissomia – Alteração cromossômica numérica que surge em consequência de três cópias de um mesmo cromossomo, em vez das duas previstas

■ U

Unidade de replicação (réplicon) – Segmento de DNA que está sendo duplicado

Unidade de transcrição – Segmento de DNA que é transcrito em uma molécula primária de RNA e suas regiões reguladoras

■ V

Vacina – Substância obtida de um agente infeccioso que estimula a resposta imunológica direcionada a este, não produzindo a doença

Valor adaptativo (W) – Medida que expressa a contribuição de um organismo ao conjunto gênico da próxima geração (sucesso na transmissão gênica) em relação a outros organismos de sua população. Mede o sucesso reprodutivo relativo que determinado genótipo confere ao organismo

Variação dicotômica – O mesmo que variação fenotípica descontínua ou binária

Variação fenotípica contínua – Tipo de variação observada em características quantitativas, na qual se verifica um contínuo de fenótipos entre os indivíduos, não sendo observadas classes fenotípicas nitidamente distintas, como ocorre na variação descontínua

Variação fenotípica descontínua – Padrão de variabilidade relativa a uma característica cujos fenótipos se apresentam em categorias nitidamente distintas e independentes

Variância ambiental – Porção da variabilidade fenotípica de um traço multifatorial decorrente das variações ambientais

Variância fenotípica total – Variabilidade fenotípica referente a um traço específico observado entre organismos de uma população

Variância genética aditiva – Fração da variância fenotípica total de uma característica multifatorial que decorre dos efeitos aditivos dos diferentes poligenes

Variância genotípica – Porção da variabilidade fenotípica de um traço multifatorial decorrente das diferenças genéticas entre os organismos

Vetores de clonagem – Plasmídios contendo em sua estrutura um fragmento de DNA de origem diferente e que foi inserido com o objetivo de ser introduzido e replicado em bactéria hospedeira

Vetores de expressão – Plasmídios que têm, além das estruturas de um vetor de clonagem, uma região promotora que irá dirigir a expressão de gene inserido

■ Z

Zigoto – Célula diploide formada pela união de dois gametas haploides (ovócito e espermatozoide) a partir da qual o embrião se desenvolve

Índice Alfabético

A

B

C

■ D

■ E

■ F

■ G

■ H

■ I

■ K

■ L

■ M

■ N

■ O

■ P

■ Q

■ R

SANTUÁRIO — 2013/1